

POLICIES AND PRACTICES IN PUBLIC PERSONNEL  
ADMINISTRATION

*Report of the Committee on  
Oral Tests in Public Personnel Selection  
of the  
Civil Service Assembly of the United States and Canada*

**ORAL TESTS IN  
PUBLIC PERSONNEL SELECTION**

*Committee on*  
ORAL TESTS IN PUBLIC PERSONNEL SELECTION  
*of the*  
CIVIL SERVICE ASSEMBLY  
*of the United States and Canada*

*Chairman:* SAMUEL H. ORDWAY, JR.  
Formerly, United States Civil Service Commissioner

CLIFFORD N. AMSDEN  
Los Angeles County Civil Service  
Commission

JAMES C. O'BRIEN  
Formerly, National Roster of Scientific  
and Specialized Personnel

WALTER V. BINGHAM  
War Department

WILLARD E. PARKER  
Louisiana Department of State Civil  
Service

WILLIAM A. HANNIG  
New York City Board of Examiners

CYRUS C. PERRY  
National Civil Service Reform League

DAVID V. JENNINGS  
Milwaukee County Civil Service  
Commission

RAY T. SLATER  
New York State Department of Civil  
Service

PAUL J. KERN  
New York City

GEORGE SLOVER  
Tennessee Valley Authority

ARTHUR W. KORNHAUSER  
University of Chicago

LOUIS J. KROEGER  
Office of Price Administration

FRED W. SMITH  
Detroit, Michigan

**ORAL TESTS IN  
PUBLIC PERSONNEL  
SELECTION**

*A Report Submitted to the*  
**CIVIL SERVICE ASSEMBLY**

*By the Committee on*  
*Oral Tests in Public Personnel Selection*

**SAMUEL H. ORDWAY, JR.,** *Chairman*

**CIVIL SERVICE ASSEMBLY**  
**OF THE UNITED STATES AND CANADA**

**CHICAGO: 1943**

*Copyright, 1943, by the  
Civil Service Assembly of the United States and Canada  
Reprinted April, 1948*



## Foreword

THE management of governmental affairs has become increasingly important as the activities of governments have grown in magnitude and broadened in scope, particularly during the last quarter of a century. This has led inevitably to an intensification of interest in problems relating to personnel administration—a function that is an essential and integral part of over-all management. The rapid extension and improvement of merit systems in national, state, and local governments and the renewed interest of many important groups in public personnel problems have marked this development during the last few years. All interested groups, including public personnel workers themselves, have long felt the serious need for a searching review and appraisal of existing personnel policies and practices and the formulation of proposals for the more complete and satisfactory performance of personnel activities. Plans for meeting this need were approved by the Executive Council of the Civil Service Assembly in 1937. The present report is one of a series which, when completed, will bring together for the first time a set of authoritative and forward-looking volumes dealing with the major phases of public personnel administration.

More than sixty outstanding personnel officials, general administrators, technical and research workers, educators, and representatives of civic, professional, and employee groups actively participated in the preliminary planning of this huge undertaking. It was agreed that the final findings and reports resulting from this comprehensive effort would be based upon special field studies of public personnel policies and practices, which would be supplemented by information obtained from existing studies and reports dealing with personnel problems and by the ideas and suggestions of those who were in a position to make helpful contributions because of their training or experience. It was further agreed that the reports should not be the work

of one person, or of a small group of persons, professing omniscience in the field. It was felt that the final reports should be the product of group effort and group thinking, which could be realized through the appointment of a series of committees whose members would give their time, energies, and ideas to make the undertaking successful.

To collect and appraise facts regarding present public personnel policies and practices, a specially recruited staff conducted field studies covering twenty-two different public personnel agencies selected because of their differences in size, location, and problems. In each jurisdiction one or more members of this field staff conducted intensive interviews with personnel administrators, technicians, departmental administrators and supervisors, political leaders, and representatives of organized employee associations. Approximately four hundred persons were interviewed during the course of the field studies. Complete notes were made of these interviews. Information and suggestions obtained in this way were supplemented by a careful study and review of other materials, such as: appropriate legislation; annual and special reports of the personnel agencies; special studies and memoranda regarding the work of the personnel agencies which had been prepared by outside organizations and disinterested persons; personnel tests, forms, records, statistics, and methods; and finally, actual observations of the agencies' operations. On many occasions, significant and helpful documentary material not ordinarily available to a researcher was placed at the disposal of the field staff.

As a result of this work, a detailed case history was prepared for each agency covered by the survey. Each case history included statements of fact regarding the personnel policies and practices of the agency; ideas and suggestions obtained from those interviewed, from reports, and from other sources; a critical appraisal of the policies and practices of the agency as they were actually working; and suggestions for changes and improvements which originated with those interviewed or members of the survey staff. The information and suggestions included in each case history were gathered and organized according to a

prearranged plan that made it possible to classify the material into broad categories corresponding to certain major aspects of public personnel administration.

The case histories and all other available materials were then placed in the hands of a number of committees for use as basic information in the preparation of final reports. Each committee was given the task of preparing a report dealing with a specified functional subject in the field of public personnel administration. In addition to the case histories, the committees were given access to supplementary descriptive and interpretative material regarding many agencies not covered by the field survey; special reports and theses relating to the work of personnel agencies and to technical and administrative problems in personnel administration; selected bibliographies; and other materials brought together by the Assembly's Headquarters Office in connection with its regular activities as a clearing house in the field of public personnel administration. Committee chairmen and members were encouraged to augment this material by consulting with persons and groups who were in a position to make substantial contributions of ideas and facts and by conducting special investigations and researches. Through the occasional issuance of memoranda and special notes, all committee members were kept currently informed of the progress being made and the problems being faced by participants in the undertaking.

Following a procedure approved by the Assembly's Executive Council regarding each committee, one person was appointed by the President of the Assembly to serve as chairman of an advisory committee to prepare an appropriate report on the particular subject or phase of public personnel administration assigned to it. The members of each committee were chosen because of their interest in, and knowledge of, the matter falling within the committee's general jurisdiction, and because of the diverse viewpoints which they could bring to the committee's work. More than three hundred persons have served on the several committees. About one-half of them are persons actively engaged in public personnel administration. The other half are general administrators, educators, industrial personnel workers,

and leading members or representatives of civic and professional groups, governmental research associations, and employee organizations.

It has been the responsibility of each chairman to initiate and coordinate the activities of his committee. The work methods of the several chairmen have naturally differed. Some have used their committees largely as sounding boards on various problems or proposals referred to them and have asked committee members to review outlines and manuscripts and to offer their comments, criticisms, and suggestions for the purpose of injecting the influence of their composite views and experiences into the final reports. Others have asked individual committee members to assume the task of bringing together all necessary material regarding a particular segment of the committee's assignment and to prepare a corresponding section of the final report. It has been the chairman's responsibility, without obtaining formal action by the committee, to reconcile differences of approach and to mold his own ideas and materials and those submitted by his committee members into a final integrated report. This procedure has made it possible for each report to represent the collective contributions of outstanding thinkers and doers in public personnel administration, and of persons engaged in other fields who have the vision, imagination, and freedom from professional introversion to propound the broad principles and objectives that should determine the role of personnel administration in the over-all scheme of public affairs.

At the very inception of the undertaking, it was stressed that each report should represent a synthesis of the most effective and desirable policies and practices on a particular phase of public personnel administration. It was contemplated that each report would be more than a mere tallying of existing practices and malpractices, and that it would thus be qualitative rather than quantitative. It was also agreed that each report would not only carry the story of the effective steps which had been taken by personnel agencies to reach certain objectives, but would go further and project beyond present policies and practices to more desirable or acceptable ones.

Any statement regarding the undertaking would be incomplete without acknowledgment of the contributions which have been received from a number of individuals and organizations. Members of the Assembly's Headquarters Office staff, past and present, have made substantial contributions and rendered effective help throughout the various phases of the project. The broad general outline of the study as a whole was first conceived by G. Lyle Belsley, then Director of the Headquarters Office. He, together with Henry F. Hubbard and Maxwell A. DeVoe, laid the foundation for the study and devised the procedures for its execution. Mr. DeVoe was responsible for the immediate supervision of the special field staff originally engaged in gathering information contained in the case studies and for coordinating the efforts of the several committees participating in the preparation of final reports. John Steven and Doris Haney Jones assisted ably in the gathering of the material contained in the various case studies. With the departure of Mr. DeVoe from the staff, the work of coordinating committee activities and the editorial work involved in preparing the various reports for publication have been the responsibility of Jeremiah Donovan.

Sincere appreciation is hereby tendered to the many public-spirited persons who, without compensation, took an active and helpful part in planning this undertaking and bringing it to its present stage of completion. Great help and much useful information have been made available to the Assembly in connection with this effort by the associations of public agencies and public officials located in the same building as the Assembly's Headquarters Office at 1313 East 60th Street, Chicago, as well as by various other organizations throughout the country. Acknowledgment is made for the assistance rendered by Public Administration Service throughout the process of printing and publishing this report. The Assembly is grateful to the several copyright holders who have permitted the quotations of copyrighted materials in the report. Finally, the entire undertaking was made possible through the finances generously provided by the Spelman Fund of New York. Without this assistance, it would be impossible for the undertaking to achieve the intended goals.

Committee chairmen and members have been able to take only a limited amount of time away from their regular activities for the purpose of carrying forward the undertaking. This fact, together with unforeseeable difficulties encountered by various committees, has made it impossible for all reports of the series to be finished simultaneously or in schematic order. It has therefore been decided to publish the reports, for the most part, in the order of their completion.

This report, *Oral Tests in Public Personnel Selection*, like all others in the series, is a document which a committee has prepared and submitted to the Civil Service Assembly. The information and recommendations presented in its pages represent the collective thinking of the chairman and his committee. It is to be stressed, however, that the report was not prepared with a view toward official approval or formal adoption by the Civil Service Assembly, its Executive Council, or its Headquarters staff, and no action of this nature is contemplated. The report is, however, as forward-looking and authoritative as an able chairman and a group of competent associates have been able to make it.

JAMES M. MITCHELL, *Director*  
Civil Service Assembly

## Preface

THIS report is a product of the combined studies and criticisms of individual members of the Committee on Oral Tests in Public Personnel Selection of the Civil Service Assembly.

Several members of the committee prepared and submitted material that served as a foundation for the report. Louis J. Kroeger presented material on the uses of the interview as an oral test. A section on planning for oral tests was prepared by William A. Hannig. The selection of interviewers and the details of testing procedure were presented by Willard E. Parker. Walter V. Bingham wrote the chapter on the training of interviewers. The troublesome subject of rating was treated by James C. O'Brien, who also assisted the chairman in writing the section on underlying principles. The chairman is responsible for the passages on "variable weights for factors," which were suggested to him by Arthur W. Kornhauser. Messrs. Paul J. Kern and Cyrus C. Perry submitted material for the chapter on legal sanctions and the limitations of oral tests. Other sections were drafted by the chairman for the comment or revision of David V. Jennings, Fred W. Smith, Clifford N. Amsden, and Ray T. Slater.

All manuscripts submitted had the benefit of intensive and detailed criticism from each of the contributors and also from Arthur W. Kornhauser, Norman J. Powell, and George Slover. A review of the original studies revealed some duplication of content and a conflict in approach which necessitated reconciliation and coordination. The materials contributed were consolidated and rewritten. After a number of conferences with the chairman, Mr. Bingham undertook a general revision of the manuscript. Subsequently, Mr. Kroeger reviewed the revised document and contributed further to its content.

The original content of the report was the work of many hands. The chairman regrets the conglomerate form that has

resulted and accepts all responsibility which he cannot fairly dodge! There is here also much of wisdom and value for which all members of the committee may properly share together a reasonable glow of pride. Comments made on textual material are carried as footnotes.

G. Lyle Belsley, James M. Mitchell, and Jeremiah J. Donovan of the staff of the Headquarters Office of the Civil Service Assembly were most helpful in assisting with the collection of material, in verifying references, and in offering criticism.

SAMUEL H. ORDWAY, JR.

New York City



# Contents

FOREWORD . . . . .	v
PREFACE . . . . .	xi

## *Chapter I. THE SCOPE OF THE REPORT*

PRINCIPAL AREAS OF COVERAGE . . . . .	1
POINTS OF EMPHASIS . . . . .	2

## *Chapter II. USES OF INTERVIEWS AND ORAL TESTS*

ATTRACTING APPLICANTS . . . . .	6
ESTABLISHING GOOD WILL . . . . .	6
The Use of Citizen Interviewers . . . . .	7
NEED FOR THE INTERVIEW IN THE SELECTIVE PROCESS . . . . .	8
Work Habits and Attitudes . . . . .	8
Social Adaptability . . . . .	9
Qualitative Evaluation of Experience . . . . .	9
Summary of Considerations . . . . .	10
SELECTION FOR POSITIONS EXEMPT FROM EXAMINATION . . . . .	11
NONCOMPETITIVE SELECTION: THE QUALIFYING INTERVIEW . . . . .	12
THE CHARACTER INTERVIEW IN INVESTIGATIONS . . . . .	13
RATING QUALIFICATIONS COMPETITIVELY . . . . .	13
Education and Experience . . . . .	14
Intelligence . . . . .	16
Adaptability . . . . .	17
Personality and Other Factors of Personal Capacity . . . . .	17

## *Chapter III. CONSIDERATIONS AFFECTING THEIR USE*

THE LABOR MARKET . . . . .	19
LOCAL PUBLIC OPINION . . . . .	21
ALTERNATIVE PROCESSES . . . . .	22
IS THE INTERVIEW THE BEST METHOD FOR THE PURPOSE? . . . . .	23
What is to be Measured? . . . . .	23
Overcoming Popular Objections . . . . .	24
Loss of Anonymity . . . . .	24
The Element of Time . . . . .	25
The Element of Distance . . . . .	26
The Element of Cost . . . . .	27

THE TYPE OF POSITION AS A GOVERNING FACTOR . . . . .	27
Supervisory Positions . . . . .	28
Administrative and Executive Positions . . . . .	28
Professional and Technical Positions . . . . .	28
Law Enforcement Positions . . . . .	29
Social Welfare Positions . . . . .	29
Labor Classes . . . . .	30
Clerical Classes . . . . .	30
USE OF THE PROBATIONARY PERIOD AS A SUBSTITUTE . . . . .	31
THE EXAMINER'S RESPONSIBILITY SUMMARIZED . . . . .	33

#### *Chapter IV. PRINCIPLES OF SOUND COMPETITIVE ORAL TESTING*

NATURE OF THE PROCESS . . . . .	35
PRE-SELECTION OF FACTORS . . . . .	36
Job Analysis . . . . .	36
Breakdown of Factors . . . . .	36
Ratability of Factors . . . . .	37
Probative Value . . . . .	37
Can Factors be Evidenced Competently? . . . . .	39
THE NATURE OF EVIDENCE . . . . .	40
Analogy of the Oral Test . . . . .	40
Relevance . . . . .	41
Materiality . . . . .	43
Trustworthiness . . . . .	44
Objectivity . . . . .	46
KINDS OF EVIDENCE . . . . .	46
WEIGHTING FACTORS . . . . .	47
Fixed Weights . . . . .	48
Variable Weights . . . . .	49
METHODS OF DEVELOPING EVIDENCE . . . . .	51
Observing the Applicant and his Controlled Reaction . . . . .	51
Eliciting Facts About Past Demonstration of Capacities . . . . .	52
Eliciting Evidence of Lack of Capacities . . . . .	54
Cautions . . . . .	55
Recognizing Symptoms of Predictable Behavior Patterns . . . . .	58
BASES OF RATING . . . . .	60
Standards for the Analysis of Evidence . . . . .	60
Standards for Applying Variable Weights to Factors . . . . .	64

Standards of Rating . . . . .	65
The Rating Scale . . . . .	67
ESTABLISHMENT OF THE RECORD . . . . .	68

### *Chapter V. ADMINISTRATIVE ASPECTS*

PLANNING FOR AN ORAL TEST . . . . .	71
Pre-Selection of Factors . . . . .	73
Planning the Rating Scheme . . . . .	74
Preparation of Oral Test Material . . . . .	75
Scheduling the Tests . . . . .	77
Plans for Proctoring . . . . .	78
Physical Preparedness . . . . .	78
QUALIFICATIONS OF INTERVIEWERS . . . . .	79
Need for Job Knowledge . . . . .	80
Mental and Emotional Qualifications . . . . .	80
SELECTION OF INTERVIEWERS . . . . .	81
Need for a Trained Chairman . . . . .	82
Departmental Representatives . . . . .	83
Use of Outside Experts . . . . .	83
Panels . . . . .	84
Conclusions . . . . .	84

### *Chapter VI. THE TRAINING OF INTERVIEWERS*

NEED FOR PLANNED TRAINING . . . . .	86
Training the Novice . . . . .	87
INSTRUCTION MANUALS . . . . .	88
Instruction in the Total Selection Process . . . . .	94
TRAINING IN THE TECHNIQUES OF ORAL TESTING . . . . .	97
The Need for Objectivity . . . . .	97
Instruction in Rating Procedure . . . . .	99
Agreement on General Policy and Procedure . . . . .	100
Value of Practice Interviews . . . . .	101
Mechanical Training Aids . . . . .	102
Some Common Mistakes . . . . .	102
TRAINING THE PROFESSIONAL INTERVIEWER . . . . .	104
Scope of Professional Training . . . . .	104
Assignments for Experience . . . . .	105
ILLUSTRATIONS OF TRAINING PROGRAMS . . . . .	105
Field Investigators . . . . .	105

Initial Assignments . . . . .	106
Training Conferences . . . . .	106
Other Programs . . . . .	107
Self-Training and Research . . . . .	108
Research as a Training Aid . . . . .	109

## Chapter VII. ORAL TEST PROCEDURES

THE EXAMINER'S MANNER AND BEARING . . . . .	110
Identification of Candidates . . . . .	111
CONDUCTING THE INTERVIEW . . . . .	111
Questions and Discourse . . . . .	111
Notes . . . . .	112
Timing the Interview . . . . .	113
Termination . . . . .	113
HAZARDS AND PITFALLS TO BE AVOIDED . . . . .	114
Pseudo-Scientific Methods . . . . .	114
Conditioned Reactions and "Halo" . . . . .	115
The Habit of Generalizing . . . . .	117
The Use of Suggestion . . . . .	118
Wishful Thinking . . . . .	119
The Struggle to Dominate . . . . .	120
Fatigue . . . . .	120

## Chapter VIII. RATING AND RECORDING

DEVELOPING THE RATING PROCEDURE . . . . .	121
Factors and Over-All Rating . . . . .	122
Internal Consistency of Ratings . . . . .	123
Independent Rating vs. Consultations . . . . .	124
Time of Rating . . . . .	125
Rating by Amateurs . . . . .	125
Predetermination of Weighting Standards . . . . .	126
Machine Averaging . . . . .	126
THE ORAL TEST RECORD . . . . .	126
Kinds of Records . . . . .	127
Longhand Notes . . . . .	127
Stenographic or Stenotype Records . . . . .	127
Sound Recordings . . . . .	127
Sound Motion Picture . . . . .	128
Recording Observations . . . . .	128

Use of the Record for Review . . . . .	129
The Danger of Excessive Formalism . . . . .	130

### Chapter IX. APPEAL PROCEDURE

HEARING OF APPEALS . . . . .	134
The Chief of the Examining Division . . . . .	134
The Commission or Administrator . . . . .	135
The Board of Appeals . . . . .	135
The Courts or an Independent Tribunal . . . . .	135
TYPES OF ERROR . . . . .	136
Evidence Overlooked . . . . .	136
Mistake in Rating . . . . .	137
Disclosed Identities . . . . .	137
Lack of Standards . . . . .	138
Irrelevant Objectives . . . . .	138
Favoritism . . . . .	138
REMEDY FOR ERROR . . . . .	139
Re-Rating Applicant . . . . .	139
Re-Rating All Candidates . . . . .	139
New Examination . . . . .	140
SUMMARY . . . . .	140

### Chapter X. LEGAL SANCTIONS AND LIMITATIONS

SCOPE OF THE DISCUSSION . . . . .	141
POWER OF THE COURT ON REVIEW . . . . .	142
STATUTORY PROVISIONS . . . . .	143
THE REQUIREMENT OF AN OBJECTIVE STANDARD OR MEASURE . . . . .	144
The Application of the Objective Standard or Measure . . . . .	145
The Determination of Qualities to be Tested . . . . .	146
The Notice or Announcement . . . . .	147
Test Materials . . . . .	149
The Standard of Rating . . . . .	151
THE RECORD FOR REVIEW . . . . .	151

### Chapter XI. EVALUATION OF THE ORAL PROCESS

CORRELATION WITH GRADES IN OTHER SUBJECTS . . . . .	154
CORRELATION WITH PERFORMANCE ON THE JOB AND PROMOTION RECORDS . . . . .	155

STUDIES OF VALIDITY OF JUDGMENT RATINGS . . . . .	156
THE PROBLEMS AHEAD . . . . .	158
Need for Rational Appraisal Standards . . . . .	159
Need for a Realistic Attitude . . . . .	160

## Appendices

A. BIBLIOGRAPHY . . . . .	163
B. RATING FORMS	
Figure 1. Rating Form (Alabama State Personnel Department) . . . . .	168
Figure 2. Rating Form (Developed and copyrighted by International Business Machines Corporation) . . . . .	170
Figure 3. Rating Form (Los Angeles City Civil Service Commission) . . . . .	172
Figure 4. Rating Form (Los Angeles County Civil Service Commission) . . . . .	173
Figure 5. Rating Form (Municipal Service Bureau, New York State Department of Civil Service) . . . . .	174

## *Chapter I*

### **The Scope of the Report**

THE purposes, techniques, and uses of interviews as described in this report. Interviews vary greatly in scope and purpose. This report refers to the entire range from over-all comprehensive oral examination down to a qualifying, non-competitive, "personal appraisal" of applicants. Principal attention, however, is given to the competitive oral testing of personal traits, characteristics, capacities, and behavior patterns that are believed to indicate the probability of success on the job.

In common practice the word "examination" is often used loosely as interchangeable with "test" in referring to a single part of a complete process, i. e., "written examination," or "oral examination." In this report the word "examination" will be used to mean the whole selective process; the word "test," to mean a single part thereof; and the phrase "oral test" to mean a measurement of certain qualifications of applicants, brought out in conference using controlled techniques.

#### **PRINCIPAL AREAS OF COVERAGE**

This survey describes current practices and states the general principles which, in the opinion of the committee, underlie sound competitive oral testing. It has been said that principle and practice are seldom combined in oral testing. This report seeks to show how to bring the two together.

The purposes, uses, and limitations of oral tests and interviews are outlined in the opening chapters. Later chapters are devoted to a more detailed description of the prevailing practices in administering them. The following questions suggest some of the topics discussed: How are oral tests organized? What is their scope? What are the qualifications of good inter-

viewers? How are interviewers selected? How are they introduced to their duties and trained to perform them? How is the test planned? What are the best methods of meeting the candidate, conferring with him, observing him, and drawing out the evidence of his fitness? How is this evidence evaluated? With what objectively defined norms is the evidence compared? How do the interviewers record their findings?

In addition to describing current procedures the report also pictures the difficulties sometimes encountered in oral testing. Such pitfalls as bias, subjectivity of judgment, misdirected questioning, and unwarranted generalization are pointed out so that they may be avoided.

One chapter analyzes certain principles considered fundamental to sound competitive interviewing. Such principles stem from basic rules governing the hiring of employees, and includes the breakdown of duties and qualifications, the selection of factors, the development and evaluation of evidence, and the definition of standards for anticipation of future achievement from the record of past behavior. A chapter on the importance of appellate procedures also discusses the variety of practices in the handling of appeals from oral test ratings. Existing legal sanctions and limitations on oral testing are summarized. The final chapter states the problems of validation that now press for further study and experiment.

Although procedures vary widely in detail among jurisdictions, it is easier to see common patterns of practice than it is to reconcile the divergent views of the authorities regarding the essential nature, purposes, and limitations of oral testing. Throughout the report attention is deliberately drawn to points of disagreement among the authors, thereby indicating the need for further research and experiment. The object of the entire discussion is to reveal an instrument of personnel selection that, when properly administered, is fair to competitors and valuable to employing agencies.

#### POINTS OF EMPHASIS

Throughout the report, and particularly in Chapter IV, the committee has voiced the need for advance planning and an



objective evaluation of the evidence brought forth during the oral test. To some readers this may appear to be an undue emphasis on the formalization of procedures and a legalistic treatment of subject matter. Under some circumstances such a criticism might be justified. However, the committee, in studying the subject of oral tests as a part of public employment processes, has been influenced by one dominant impression—the lack of appreciation on the part of many in the public personnel field of the need for standards and self-discipline to make the interview a scientific tool in competitive testing.

This lack of appreciation is understandable, though not excusable. Rare is the man who does not harbor the conviction that he is a good judge of human nature and human potentialities. This talent, he feels, consists in bringing to bear on a particular situation his interpretation of past experiences with his fellowman; by this process he may arrive at a conclusion on the basis of memories to which his observations of the individual give rise. Convinced of his own acumen and sincerity, he rarely stops to think that his right to exercise this dubious assurance and to act on subjective conclusions may legitimately be challenged.

But one area in which such conclusions may be challenged is in the field of public employment under a democratic merit system. Here competitors for employment have the implied right—reinforced more often than not by specific legal mandates—to match their personal qualifications against one another in relation to predetermined objective standards. They should all start from scratch, run the same course, and be timed by the same stop-watch. If significant variations occur, if the hurdles are higher for some competitors than for others, equality of opportunity vanishes. Moreover, the damage thus done is double: some competitors are handicapped unduly, and the spurious results of such tests mislead the prospective employer as to the relative worth of applicants.

All of these circumstances have combined to make the oral test a bone of contention. Lack of appreciation of the considerations involved on the part of the personnel agency has often given rise to lack of planning, absence of standards, and

significant variations in continuity of procedure. Competitors are quick to sense these weaknesses and often seize the opportunity to condemn the tool as well as the workman who has misused it.

It is for these reasons, then, that the committee has placed considerable stress on the need for planning, for standards, for consistency of procedure, and for objectivity in appraising evidence. This emphasis may seem overstressed to some, but it is an emphasis which, it is hoped, will serve to focus attention on an important area of the selection process which has been too often neglected.

## *Chapter II*

### **Uses of Interviews and Oral Tests**

**T**HE interview, as used in public personnel selection, takes many forms and serves a variety of purposes. In some instances it constitutes a test; in others it does not. A personal discussion with a prospective applicant concerning opportunities in the public service is an interview, but is not a test. An interview between an eligible and an appointing officer is seldom a test in the sense that predetermined standards are employed and evaluation of evidence takes place. While this report recognizes the roles of interviews such as these in the overall process of selection, they are distinguished from the true oral test for purposes of subsequent treatment.

It is also recognized that, whatever the form or ostensible purpose of the interview may be, it serves one necessary end—to humanize an otherwise bureaucratic relationship. To the applicant, participation in a civil service examination is a highly personal experience, whereas the examiner is prone to view it as a matter of impersonal, objective routine. The interview presents the one opportunity along the way to personalize the process. Moreover, it often presents an opportunity for engendering the goodwill which springs from letting the candidate know that he is being considered as a human entity, rather than as an array of skills, talents, and similar abstractions.

Interviews often do meet the criteria of a valid competitive or noncompetitive oral test of factors relating to fitness for a position. As such, they may take the form of a supplementary evaluation of knowledge or experience already measured in other parts of the examination. Or they may evaluate evidence of the possession of personal characteristics, traits, and capacities impossible to assess in other types of tests.

The personal interview may be used for other purposes not pertaining to formal selection, such as in employee supervision,

in counseling, in the adjustment of grievances, and in exit from the service. However, these topics are to be described in other reports in this series. Consequently the principal purpose of this report is to consider the interview as part of the selective process.

### ATTRACTING APPLICANTS

Enlightened personnel administration recognizes the importance of seeking out well qualified candidates for public service, particularly for high administrative, technical, and professional positions.<sup>1</sup> It is no longer sufficient merely to examine those who are out of work. Advanced public personnel agencies now publicize the advantages of the service, and, taking a leaf from industrial practice, they frequently send representatives on recruiting missions to colleges and professional schools, and to the meetings of professional and trade groups. These representatives usually describe the advantages of public service, the qualifications required, and the methods of qualifying. They frequently urge interested persons to confer with them individually to discuss specific opportunities. Sometimes these interviewers even search out qualified potential applicants and through personal discussion persuade them to apply for forthcoming examinations.

While not part of the formal process of selection, these interviews are important to its success. Unless well-qualified candidates can be found, no competitive test can adequately fulfill its function of supplying a list of competent eligibles.

### ESTABLISHING GOOD WILL

The technical aspects of civil service examinations often seem incomprehensible to the public. The use of mechanical and impersonal tests may make the applicant feel that some of his qualifications for the job are not being thoroughly considered. He wants an opportunity to talk personally with a representative of the examining agency. Some jurisdictions have met

<sup>1</sup> See Civil Service Assembly committee report, *Recruiting Applicants for the Public Service* (Chicago, 1942).

this need by making it possible for any applicant to discuss his application with an examiner.

This type of interview is intended, not to draw out ratable evidence, but to personalize the employment relationship between the public and the government. This is a costly and time-consuming process, but the good will it creates fully compensates for the expense.

When the number of applications is small, and all applicants can be interviewed, such an interview may be used to observe personal characteristics for rating purposes; but this makes the interview part of the selective test, dealt with hereafter, and should not be confused with conferences having the sole purpose of enhancing good will.

### *The Use of Citizen Interviewers*

Some jurisdictions create good will for the merit system by inviting prominent persons to serve as members of oral examining committees. An active jurisdiction, in the course of a year, may in this way gain the interest and cooperation of a large number of representatives of the professional, administrative, and civic life of the community. Even though they are not always asked to rate the candidates, these citizens, through service on such panels, learn at first hand the importance of public employment testing and the value of the civil service system to the public.<sup>2</sup>

When the outside participants are experts in the field for which the test is held, they are usually asked to help in rating candidates. This process tends to increase their confidence in the tests; but, unless the participants are given systematic preliminary training, it may add a subjective quality to the ratings. A few jurisdictions, instead of undertaking to train these outside participants, have resorted to the questionable procedure of inviting them to submit their ratings, and then, in computing and publishing the ratings, have disregarded the conclusions of the outsiders.

<sup>2</sup> See Alice Campbell Klein, *Civil Service in Public Welfare* (New York: Russell Sage Foundation, 1940), Chapter 10; and Marie Dresden Lane, "The Education of an Examiner," *Survey Midmonthly*, April 1938, pp. 105-07.

On the whole, it may be said that, provided a sound method of testing is used, participation by outsiders in the oral testing process increases good will for the system. The public learns that the tests are honestly conducted, free of political manipulation, and that the selective process is not as cold and impersonal, or as slipshod and ineffective, as is sometimes alleged. The good will thus earned may bear fruit in many ways, such as helping the personnel agency to obtain adequate funds and support from legislative and executive branches of the government.

#### NEED FOR THE INTERVIEW IN THE SELECTIVE PROCESS

As one part of the process of comparing the merits of applicants, the oral test seeks to appraise personal factors that may vitally affect performance on the job. Every employer is concerned with elements of personality and character that are less easily recorded and weighed than education, previous employment, and demonstrated skill.

#### *Work Habits and Attitudes*

The employer would like to know not only that the applicant *can* do the work required but that he *will* do it. What is the applicant's attitude toward his work? What are his personal interests? Many an able employee has devoted great ingenuity and effort, at the expense of his job, to activities that interest him more than his job. He may be brilliant, but at the same time lazy or indifferent. Emotional difficulties may prevent him from applying his full ability to his job. In short, the attitudes, interests, temperament, and traits of character of a candidate may be factors of fitness quite as vital as his skills and knowledge. He may be able to turn out an enormous quantity of accurate work under certain kinds of pressure, and yet, under routine conditions, be habitually inaccurate and dilatory. It is desirable, then, to find out not only what he can do, but also what he may be expected to do under given conditions.

Again it may be important for an employer to know about an applicant's aptitude for learning. How quickly can he learn

new duties and meet new problems? Will he become impatient for promotion, or, however expert, will he prove incapable of assuming additional or new responsibilities? These questions can be answered not by proof of skill, prior training, or present ability so much as by proof of past behavior, of aptitudes, ambitions, and emotional responses.

### *Social Adaptability*

Another question of primary importance to the employer is, Will the candidate fit into my organization? Every enterprise, every office, every crew of workers is a social system in itself. The effectiveness with which its basic purpose is fulfilled depends in no small part on the ability of its chief and his associates to maintain social equilibrium within the organization and also within the larger systems of social relationships which involve contact with the public, customers, or collaborating agencies.<sup>3</sup> A worker, a supervisor, an administrative aide, or a chemist who irritates his associates, annoys his superiors, or otherwise fails to function as an integral part of his organization, may seriously hamper the enterprise even though he completes the specific duties of his position with conspicuous ability. It is little wonder that employers make every effort to learn the answer to the question, Will the candidate fit in? In considering the fitness of an applicant, then, his behavior, emotional adjustment, and aptitude assume an importance comparable with his knowledge, skill, and training.

### *Qualitative Evaluation of Experience*

Behavior patterns are not usually disclosed in a written test, nor are they reflected in a quantitative record of experience. For that reason examining agencies have increasingly turned to the oral test as a means of completing the record of an applicant's suitability. Most examiners agree that for testing attitude, adaptability, and personal capacity, no other tech-

<sup>3</sup> See F. J. Roethlisberger and W. J. Dickson, *Management and the Worker* (Cambridge: Harvard University Press, 1939), and Chester I. Barnard, *The Functions of the Executive* (Cambridge: Harvard University Press, 1939).

nique can take the place of the oral test. Many believe also that the technique of the controlled interview or oral test provides the only effective way to find out the true value of a person's prior experience. Examiners have long recognized that "exposure" to education, and the time "occupied" in relevant employment are not conclusive evidences of competence. Employers want to know which of the abilities needed on the job a candidate now possesses—acquired, they care not how—from the sum of his life experience. The interview usually serves to draw out verifiable evidence of *demonstrated* capacities. It is at this point that personality and experience tests dovetail. The probability that a candidate can gain cooperation or accomplish objectives in the face of opposition, for instance, is partly a matter of his personality and partly a matter of his experience in such situations. Thus, the evidence of demonstrated abilities or of lack of abilities, brought out by interview, is used in some jurisdictions to supplement preliminary evaluations of a quantitative record of experience, as well as to judge additional factors of personal fitness.<sup>4</sup>

### *Summary of Considerations*

It is generally agreed that for many classes of positions which demand a high order of ability to deal personally with subordinates, associates, or the public, an interview is indispensable. When the personality of a candidate is important to performance, it is essential that someone responsible for selection interview him personally, see him in action, observe the way he conducts himself during the give-and-take of a personal conference, and afford him an opportunity to establish his ability to do the job. The interview also provides an opportunity to probe the circumstances surrounding each item of the experience appearing in the written statement of past employ-

<sup>4</sup> The California State Personnel Board, to cite one example among many, instructs the members of its Qualification Appraisal Boards to afford the candidate opportunity to go into greater detail concerning his qualifications than it is possible to record on an application form. They also use this occasion to verify statements made on the application form. The New York City and United States Civil Service Commissions do the same for certain classes of positions.



ment. A personal contact of this kind adds assurance that suitability has been fully checked prior to appointment.

In general, then, it may be said that interviews, whether formal or informal, are needed to supplement other sources of information about the candidate's abilities, potentialities, personality, and character, to the purpose of discovering whether or not he will, if appointed, fill the position as required.

The interview, however, is a costly and time-consuming technique. There is general agreement that, when other more satisfactory or objective methods of gauging any factor of fitness exist, they should be used.

#### SELECTION FOR POSITIONS EXEMPT FROM EXAMINATION

The traditional method of selecting employees in industry until recently was the personal interview, followed by a short tryout on the job. Employment interviewers have held their positions in part because they were believed to be good judges of human nature.

In its simplest form the employment interview was designed to ascertain the over-all fitness of each applicant appearing before the interviewer. The word "fitness," in this sense, embraced education, knowledge, experience, personality, character, presence, and all other factors which might be considered pertinent. It was customary to look over the data on an application blank and, after friendly conversation, arrive at a conclusive judgment, that is, general estimate of the qualifications of the applicant and of his suitability for the job. This is judgment rating, pure and simple; there is no attempt to measure specific abilities objectively, or to set up standards of comparison so that the interviewer's evaluations might be quantitatively recorded. In recent years, more progressive personnel policies are replacing the old method of selection; fallible judgment is giving way to scientific tests.

In the public service, the over-all interview has been used extensively. It is still used very frequently in filling positions outside the boundaries of the merit system, at least when the exigencies of political pressure permit any choice at all. Usually

little choice is permitted in such cases. The interviewer is most often the appointing officer in the department where the vacancy exists, and the most important consideration is proper political sponsorship.

#### NONCOMPETITIVE SELECTION: THE QUALIFYING INTERVIEW

In making appointments to noncompetitive positions, where minimum qualifications must be established without comparative rating, certain quantitative requirements of education and experience are commonly the prerequisites to eligibility. Written and oral tests may also be given. Unrestrained judgment, here, is not the sole criterion of qualification. Yet the central personnel agency is responsible only for certifying that the nominee is or is not qualified; it is not called upon to compare the merits of qualified applicants. Under these circumstances the agency sometimes uses the interview to supplement other evidences of fitness, and often relies upon the conclusion of the interviewers to furnish a final basis for the acceptance or exclusion of the applicant.

When civil service agencies first recognized the importance of objectivity in all tests they sought to extend it even to noncompetitive qualifying tests. The interview tended more and more to deal with those attributes best judged by a face-to-face conference. But not enough effort was made to classify those attributes according to the factors which could yield objective evidence. The uncontrolled judgment of the interviewer still prevailed. His sole problem was to determine whether or not the candidate had the "appearance" or "judgment" suitable for the job. He reported "yes" or "no" and on this basis the applicant was often hired or excluded from further consideration.

With the passing of time, specific factors of qualification or disqualification are being required more and more, even in noncompetitive tests. There should be no difference between a competitive and a noncompetitive oral test in so far as the factors of qualification to be measured are concerned; but, when the interviewer is not compelled to fix comparative rat-

ings as between candidates on each factor, he generally pays less attention to essential standards of objectivity and reliability.

Although the interview is thus widely used for qualifying or disqualifying applicants for noncompetitive positions, its use for this purpose, without the safeguards shortly to be described, remains suspect.

### THE CHARACTER INTERVIEW IN INVESTIGATIONS

Fitness for a position frequently depends on the character of a candidate as well as on his personality and ability, and this may necessitate use of another form of the interview. An applicant who has passed all written and oral tests for a position may be considered as qualified until the report of a personal investigation, an unfavorable voucher, or a letter referring to weakness of character or impropriety of previous conduct raises a question of possible disqualification. At this point he may be questioned about the allegation as to his character or conduct. This kind of interview is noncompetitive in the sense that the final judgment of the interviewer is not scaled to weigh the fitness of character of one applicant as against all others. Again he is either qualified or disqualified. This type of interview, considered further in Chapter VI, is actually part of the investigating process, but its validity and effectiveness are governed by the same principles as other forms of interview subsequently discussed.

### RATING QUALIFICATIONS COMPETITIVELY

When applicants are to be ranked in order of merit, the examining agency may depend exclusively upon ratings based on a test of skill or knowledge and a statement of experience. This practice is usually followed when ability to deal with people is not important on the job. In some cases a qualifying interview may be required to make sure that no candidate is certified who fails to meet minimum requirements of personal fitness. But to do a thorough competitive job the interview should be treated as an integral part of the competitive process.

Most jurisdictions use the term "oral test" when referring to

the competitive interview. Others speak of the "oral interview," although all interviews are oral by definition. The competitive interview process has been abused on occasion by stretching its use to rate at one step all possible factors of qualification, including education, experience, knowledge, intelligence, and adaptability, as well as traits or characteristics of personality which may or may not be essential factors of fitness for a particular job. This practice is not generally approved, for many of these factors may be tested better by other means.

### *Education and Experience*

The method of rating education and experience is usually based on statements prepared by applicants, setting forth in more or less detail the record of their training and employment, with such supporting documents as certificates of schooling, courses completed, marks earned, degrees received, theses or works published, and letters of recommendation or vouchers from teachers or employers. Examiners realize, however, that the courses taken, the marks earned, the work undertaken, or the promotions granted are not so important in rating the applicant as the capacities gained from studies and experiences that enable him now to perform effectively the work of the position offered. It is because quantity of experience, titles of positions held, and pay received have been found unsatisfactory as final criteria that examiners have sought to ascertain and to rate the quality of educational and work experience.

The quality or value of the experience gained in any particular position can rarely be ascertained from the applicant's own written statement. Vouchers or recommendations of previous employers often leave unsaid the most important and vital facts about experience. One method of supplementing incomplete or inadequate information is to conduct personal investigations among previous employers, but if the investigation is thorough for each applicant, it becomes costly. It may be added that sometimes an incomplete or unskillful investigation may destroy competitive fairness.

By the time an investigation is made, previous employers may have forgotten the applicant personally and records of performance kept at the time of employment may be incomplete and inadequate. It is seldom possible for an investigator to reconstruct all the significant conditions of prior experience. A more satisfactory method of supplementing experience statements is the well-planned questioning of the applicant during a personal interview. Facts about the amount and nature of work performed, the kinds of responsibilities assumed, the extent of supervision, conditions at the place of work, and advancement in responsibility may all be revealed. It may be possible to find out what kind of problems the candidate has faced and how he met them. This kind of information enables the interviewer to place a qualitative value on the candidate's experience and will afford a basis for competitive rating.

Some authorities assert that examiners should not rate facts regarding the candidate's training and experience brought to light during the interview, if such factors are also to be rated independently on the basis of data on the application form. It is said that these factors should be rated only once, that otherwise they receive undue weight in determining the final ranking. Those who advance this argument disregard the fact that education and experience are normally reflected also in the written examination mark; this is inevitable if the examination is constructed to test the applicant's general or technical knowledge as well as his alertness. Others believe there is no valid reason why one factor in the applicant's qualifications should not be measured two or three times by different methods. They urge, indeed, that the practice of measuring an important characteristic of the candidate in two or more kinds of test *should* be followed if each of the tests is known to have validity. The average of two or more measures of the same thing is more likely to provide a correct mark than is a single measure. If the measures are known to differ in validity, different relative weights may be assigned. The degree of knowledge in particular fields of work, which is usually ascertained through written or demonstration tests, may be supplemented

by the personal interview. Here again, however, the dangers of subjective rating should be weighed fully before resorting to use of the interview as a method of testing knowledge.<sup>5</sup>

### *Intelligence*

Tests of intelligence have been developed under controlled conditions in a variety of written forms. Such factors as quickness of comprehension and directness of response probably can be ascertained by written examinations more objectively than in an interview. Yet there are jurisdictions which continue to believe that the judgment of an interviewer as to the comparative intelligence of applicants is essential to an adequate selection procedure. If there is merit in this contention, it must be found in the possibility that during an interview an examiner can observe, not how intelligent the applicant really is, but whether he gives others with whom he talks the impression that he is intelligent, alert, quick to grasp a point, or dull, slow, and

<sup>5</sup> W. V. Bingham has contributed the following footnote to this subject:

"To be sure, a well constructed written examination furnishes an excellent measure of this important factor of knowledge; but when in the course of an interview the candidate discloses a deplorable lack of technical competence in an indispensable part of his field of knowledge, some essential topic which perhaps was not adequately covered in the written examination, is there any good reason why the oral examiners should be forbidden to take cognizance of such a deficiency?

"Theory has been at fault. It has started from a false assumption, namely, that each distinguishable factor contributing to a candidate's competence and suitability can be appropriately measured by means of only one kind of instrument—a written examination, or an interview, or a rating of the experience record, or a physical examination, or a character investigation. The inference has been drawn that nothing more of any significance can be learned about a designated factor of a candidate's qualifications by applying to it, quite independently, two or three of these instruments of measurement.

"Far better it is to utilize each instrument, each type of examination in so far as it is known to furnish valid measures of factors to be appraised, as a means of supplementing and correcting the findings of the other examinations."

To this, L. J. Kroeger adds the following comment: "If the written test fails to eliminate a person deficient in technical knowledge, the responsibility should be placed back on some other phase of the testing process and the interview should not, at least for the purpose of one candidate, become a test of knowledge as well as an evaluation of education, experience, and personal traits."

W. A. Hannig draws attention to the fact that a deficiency in knowledge brought out in an oral examination "may point to the need for revising the written test: (a) to make it more inclusive; and (b) to establish 'part pass marks' on segments of the written test embodying indispensable knowledge in specific areas."

uncomprehending. The competitive value of such observations is dubious.

### *Adaptability*

There are those who believe that the personal interview is the best method of determining whether or not an applicant can adapt himself to a particular condition of work, a particular type of supervision, organization, or discipline. Of course, applicants are entitled to know the conditions under which they will be expected to work, and this subject should be discussed with them in advance of the appointment. But that is a wholly different matter. Many examiners doubt whether the interview is a sound medium for testing adaptability, particularly when it is one of the characteristics for which controlled written tests may be developed.

A candidate's reaction to the conditions of the interview itself may furnish some evidence as to his adaptability. There is danger, however, of jumping to a conclusion not justified by evidence—the danger of interpreting subjectively the meaning of a candidate's behavior. On the other hand, the interviewing process may be valuable for rating adaptability if it is directed to a discussion of previous conditions of employment in which there has been a need for this trait.

Still another alternative test of adaptability is, of course, the probationary period. This working test, however, is an expensive method of testing particular traits. It ought to be reserved as a final check of the findings of the tests; it should not be used as a substitute for them.<sup>6</sup>

### *Personality and Other Factors of Personal Capacity*

Most civil service jurisdictions continue to use the oral test to rate many personal factors. Some jurisdictions maintain that many of the most important factors of personality can best be evidenced and rated separately by means of the interview. Other jurisdictions are satisfied with their evaluation of per-

<sup>6</sup> See the Committee report on the Probationary Period, a companion volume in this series.

sonality in terms of the reactions of examiners to the "person-as-a-whole," as observed during an interview. Still other jurisdictions combine these methods.

The use of the oral test as a competitive measure of personality and personal fitness presents many difficult problems to the consideration of which the main body of this report is devoted.



## *Chapter III*

### **Considerations Affecting Their Use**

**W**HETHER or not the competitive interview should be used in examining applicants for a particular position depends on certain qualifying considerations. The condition of the labor market, the number of applicants to be examined, the element of time, local public opinion, the availability of qualified examiners, the relative cost of administering written and oral tests, and the state of the budget are generally taken into account.<sup>1</sup> Under ideal conditions all useful tests should be employed; but conditions are seldom ideal. The relative effectiveness of different types of test has to be weighed along with the suitability of each type for testing the particular qualifications required for the class of position in question. Each of these considerations bears on the final decision.

#### **THE LABOR MARKET**

The number of available applicants often weighs heavily in deciding whether or not to use the oral test in a given examination. When there are barely enough candidates to fill anticipated vacancies, examining agencies sometimes forego the use of the interview, but most agencies believe it is not sound practice in the face of an adverse labor market to resort to lower standards and a relaxing of testing methods. It is rather the responsibility of the agency to make use of the interview whenever it is justified on the basis of other considerations yet to be discussed. If, by the use of reasonable standards, applied in part through the interview, a sufficient number of qualified candidates cannot be obtained, then attention should be di-

<sup>1</sup> For example, when there are only four or five competitors for a highly technical post, such as chief engineer or medical director of a great hospital, the cost of a thoroughly fair yet comprehensive objective written test may be greater than the cost of securing a board of outstandingly qualified and distinguished oral examiners. But that is only one of many considerations.

rected to the correction of the conditions that tend to limit the competition. Perhaps the salary scale for the position is out of line, or the methods of recruiting inadequate.<sup>2</sup>

Recent years have seen a constantly increasing interest in the public service and this points to the conclusion that almost any type of labor shortage may be overcome by the proper methods of recruitment. (Abnormal conditions brought about by the war are, of course, an exception.) Thus, in normal times, the more common problem faced by public personnel agencies is coping with numbers of competitors far beyond the needs of the service. This general tendency toward an excess of applicants, many of whom are adept in the written tests, tends to emphasize the importance of that form of the interview which adds to the selective process. It is sometimes alleged that the oral test is used for the indiscriminate elimination of large numbers of people by arbitrary standards—a practice which is no more justified than is the relaxation of standards in the event of a shortage of applicants.

On the other hand, some jurisdictions abandon the interview when large numbers of candidates appear for the reason that much time and money may be expended in interviewing several hundreds or thousands of candidates when available positions are numbered in tens. Under this condition, intelligent attention should be directed to positive recruiting and the sifting out of a large proportion of the less qualified applicants in those parts of the examination that deal with the candidates en masse. Resort may be had to stringent entrance requirements and to high standards in the written test, so that the number of those who ultimately reach the interview is held to reasonable proportions. Further, when the weight of the oral test is properly limited, or the evidence developed on the oral test is used to increase or decrease initial ratings on experience, it is necessary to call only a limited number of candidates who have met the preceding requirements. If the rating system used is planned so as to make possible a change in relative ranking of perhaps not more than 50 places it is nec-

<sup>2</sup> See the Civil Service Assembly committee report, *Recruiting Applicants for the Public Service* (Chicago, 1942).

essary to call only those who would have any possible chance of being reached for certification if they received maximum credit on the oral test. The other candidates, not called, are not eliminated; they are notified that they will be called as it becomes possible for them to earn a rating that is to be reached for certification.<sup>3</sup>

### LOCAL PUBLIC OPINION

The attitude of the public toward the use of oral tests is probably the most compelling consideration of expediency which the public personnel agency in a democracy must regard. It may be said that, on the whole, little adverse public reaction has been encountered in those places where the interview has been administered honestly and objectively. Whatever opposition is met usually stems from a vocal few, most of whom have been unsuccessful in their attempts to gain a place in the public service. (It is bitter medicine to be told that one lacks acceptable personality for public service.) It is, however, the responsibility of the public personnel agency to meet this kind of a belligerent and active minority tactfully in the forum of public opinion, relying on a factual presentation of the working methods of the agency and a description of its efforts to meet the needs of the jurisdiction it serves. Simply to proclaim honesty, objectivity, and purity of motive, when charged with the opposites of these virtues, only lays the agency open to the additional charge of evasion. It is essential that the purpose and details of procedures be patiently explained and pressed upon public attention, so that the people may judge for themselves the fairness of the process.

In the final analysis, public opinion in any jurisdiction is controlled largely by the success which the public personnel agency has had in convincing citizens of its integrity. The use of objective methods by the interviewing boards may always be evidenced by a complete record of the procedure and find-

<sup>3</sup> Many examination announcements of the federal government now state: "Oral examination will be given to a sufficient number of competitors to meet the needs of the service in the order of their standing. Adjustments may be made in the initial ratings in accordance with the evidence secured for the record as a result of the oral examination, and applicants who fail the oral examination will not be eligible for appointment."

ings of those boards. The goal of winning public approval is advanced in some communities by inviting leaders of the community to serve on the interviewing boards and by publicizing their participation in the selective process. But if the process itself is unsound this practice will act as a boomerang.

### ALTERNATIVE PROCESSES

There are some classes of employment in which an oral test is an essential part of the qualifying process because certain elements of capacity, personality, and character cannot otherwise be tested and are requisite to effective performance on the job. There are other classes which require, chiefly or entirely, capacities that can be measured more adequately by other means. The necessity for the particular processes should always be considered.

Some jurisdictions have undoubtedly been accused of resorting to "subjective" interviews in disregard of a requirement for "practical tests," simply because they have labeled as "personal interview" that which should have been described as a demonstration test, or trade test. For instance, when the telephone operator is called for an "oral" which actually consists in a demonstration of ability to operate a switchboard or when the tree pruner is called for an "interview" wherein the principal purpose is to observe him climbing a tree, they are competing in the most effective or practical test possible. This fact should be clearly pointed out. Appropriate processes should be appropriately labeled, and their purposes and necessity made clear to competitors.

Lack of competent test technicians does not justify including an oral test as a corrective. An oral test can never be made to compensate for the failure of examiners to do their preliminary testing well. The interview is used to measure particular factors and characteristics. It is not suitable for measuring many others. It is not designed to carry the burden of other processes when the technical practices of a jurisdiction are so faulty that more appropriate tests were not properly developed or administered.

Failure to make proper use of other kinds of tests has led in some jurisdictions to charges of manipulation, extreme subjectivity, and improper use of the oral process. A jurisdiction will never be able to remedy or perfect its written test techniques, its practical demonstrations of skill, or other qualifying processes by throwing in an "over-all oral test" and weighting it heavily to overcome shortcomings in these prior procedures. If a medical examiner is unable to determine positively that a candidate suffers a seriously disabling defect, it should not become the function of oral examiners to pass upon that question. It would be far better to dispose of the case finally on the frank basis of doubt.

#### IS THE INTERVIEW THE BEST METHOD FOR THE PURPOSE?

Resort to the competitive interview should be avoided until an honest affirmative answer can be given to the question of whether or not an oral test is the best method to achieve the purpose in a given examination. The decision to use an oral test should not be the result of a general policy either of uniformity or convenience. It is ridiculous to include an interview among the qualifying processes in every examination. Rather, it should be used only where the experienced insight of a competent examiner recognizes its need. A number of considerations will bear upon his conclusion.

#### *What Is to be Measured?*

For many classes of work, tests other than the interview are the best available means to predict relative success or failure. The elements of fitness to be tested are of the utmost importance. Achievement and aptitude tests measure important traits which some jurisdictions still seek to judge subjectively by interview. Almost without exception knowledge of subject matter can be tested more successfully by written examinations or by demonstrations of skill. Reliability or "character," so frequently judged without a great deal of foundation by interviewers, can be measured much better by a personal investigation; the use of fingerprinting may reveal a serious legal or

social offense. Where manual skill is essential, no good reason exists for substituting the oral interview for an actual demonstration of that skill. Finally, it is futile to seek to judge physical characteristics by lay observation when the medical examiner can make much more exact findings and when track and gymnasium apparatus can be resorted to for actual demonstrations of physical agility.

Even when other methods do not exist which are more objective than the interview in evaluating a given quality, it may still be true that these other methods are at least more reliable. If so, the more reliable method should be used.

This is not to say, however, that the interview should only be used when nothing else is available. There is no preconceived and stereotyped pattern of testing for any quality or for any class of employment. Consideration should be given to that combination of qualifying processes which will be most effective under all the circumstances for a given class of employment.

### *Overcoming Popular Objections*

Other qualifying processes may be more popular than the oral test. However, unless these other methods are of at least equal effectiveness, the question of popularity should have little weight. Sometimes the interview, because of an unfortunate local history, has incurred the wrath of the citizenry to an extent that to continue its use would jeopardize the entire personnel program. In this circumstance it is urged that, as an alternative to complete abandonment of the interview, it be used carefully and intelligently for the limited number of classes for which its value is indisputable. Its proper practice should be generally publicized and its necessary purpose impressed upon the public consciousness to the end that any unpopularity based on misinformation may be dispelled.

### *Loss of Anonymity*

The fact that the use of the interview destroys the anonymity of candidates, which examiners are careful to preserve in other

parts of the testing process, seems to be an important consideration in the minds of many persons. A candidate is either introduced to the examining board by name, or is soon identified from an application form containing specific facts about himself.

The preservation of anonymity, however, has probably been given more emphasis than it deserves. The procedures for safeguarding anonymity which still persist in the conduct of written tests sprang from necessity; the judgment rating of old essay type answers and theses was frequently open to abuse. Those processes, it is well recognized, permitted such wide latitude for uncontrolled judgment that it was necessary to conceal the identity of all candidates to prevent discrimination or favoritism. Similar abuse of oral testing processes, before objective methods were introduced, suggested that anonymity should be maintained at that point too. Today, under controlled conditions, both for short answer written tests and for the development of objective evidence to be recorded and rated in oral tests, anonymity is no longer as imperative as it used to be. Oral tests can be made so objective that, even if the identity of the candidate is known, unequitable scoring can be detected from the record. As objectivity in the conduct of tests is attained the importance of anonymity decreases. Should an oral process permitting free use of subjective judgment on the part of the raters be, for some reason, not amenable to corrective change, the best remedy is to secure raters who are so well known and respected in the community as to be above suspicion.

### *The Element of Time*

Time is a compelling factor in determining the extent to which the interview is used. In a jurisdiction where rules require that all candidates be tested in all parts of the examinations regardless of their success in preliminary stages, the time element assumes such proportions that oral tests have to be omitted in large examinations for many classes of employment in which they can be of great value. Under these circumstances rules or other basic authority should be amended

Time also involves the element of cost, for the agency must pay for the time taken by the staff to conduct interviews of candidates separately. Further, in an emergency, the time limit within which an eligible list may have to be produced may preclude interviewing the required number of candidates. Moreover, when the number of candidates is large, the long period over which the interviews must extend increases the difficulty of maintaining a uniform standard, and the competitive fairness of the process is thus impaired.

Under most circumstances, however, there is little excuse for avoiding oral tests because of the time element. A proper scheduling of examinations will prevent any large number of different registers from becoming exhausted at one time. When many competitors apply for a class of work in which a very few vacancies exist, other eliminators should be introduced at earlier stages in the qualifying process to remove the burden from the interviewing process. Moreover, there is increasing acceptance of the practice of summoning for oral tests only those who, if they receive full credit therein, may reasonably be within reach for certification to the vacancies anticipated. It is useless to call hundreds for interview if they can have no chance of being reached for appointment.<sup>4</sup>

### *The Element of Distance*

Distance is a consideration closely related to time and cost. In municipal jurisdictions this is a relatively unimportant fac-

<sup>4</sup>War-time experience with the National Roster of Scientific and Specialized Personnel suggests that, for many technical positions, the oral testing of a small number of applicants who possess maximum qualifying experience will become widely used in future merit system practice. Under such a procedure applications can be kept open constantly; experience records of all applicants can be thoroughly classified and coded; and the ideal or maximum (instead of minimum) requirements for vacancies can be announced by the examining agency, either individually or by classes.

Punch cards containing the coded data can be run through a sorting machine, quickly and automatically segregating the records of a small number of applicants who most nearly meet the announced ideal requirements. These applicants can then be called in for written or oral tests, or both. A procedure of this sort has several advantages: the process remains fully open and competitive; the numbers called in for testing may be kept as small as may be desirable; time and expense are saved; and no applicant is eliminated from fair consideration again for any future test for which he may possess the nearly ideal or maximum requirements.



tor, but in state and federal services it is of major consequence. When the area over which the test is extended goes beyond the limits of a single community, it becomes necessary either for the personnel agency to incur considerable cost in having the interviewing board travel from place to place to meet with the candidates, or else the expense must be transferred unequally to the various candidates, each traveling to a central point from his place of residence.

When neither of these practices can be afforded, and when the personnel agency, as the next best substitute, provides different examining boards in different parts of the area, there arises the danger of inconsistency in application of the standards by the different boards.

One solution of this problem has been to use different boards in various parts of the area, but to assure a common standard by having at least one representative of the personnel agency serve with local experts on each such board, and travel in turn to each of the places of interview. In the California state service, experience has indicated that when one member of the personnel agency's staff sits with each of these boards and serves as its chairman, there may be no greater variation in standards than when a single examining board is required to extend its deliberations over many days.<sup>5</sup>

### *The Element of Cost*

Cost, like time and distance, is a compelling though not a logical factor in deciding whether the interview shall be used. A jurisdiction ought to have sufficient funds available to provide for the use of the interview in the examining process when there is reasonable need for it.

### THE TYPE OF POSITION AS A GOVERNING FACTOR

Whether an oral test will be included in a given battery of tests depends primarily upon the type of employment involved.

<sup>5</sup> There is, of course, the alternate solution of scheduling the written test and the oral test on the same day. This is feasible only where the number of candidates involved is comparatively small, and where it is probable that only a few will be eliminated through the written test.

Some of the considerations that should guide the examiner in determining whether an oral test should be included in the examination for a particular type of position are covered in the discussion which follows.

### *Supervisory Positions*

In supervisory employment whether clerical, professional, technical, or in a skilled trade, the use of the interview as a qualifying process is of primary importance. Here, the function of the interview is to record the evidence of possession of supervisory capacities or personal traits believed to be valuable in the performance of supervisory work, and to evaluate its significance alone and in the light of known conduct and behavior patterns. When properly conducted, such an interview can be expected to indicate probable success or failure in discharging supervisory responsibilities.

### *Administrative and Executive Positions*

The duties of executives and administrators require still greater discrimination in the use of the interview in evaluating personal fitness. Behavior patterns, experience in human relationships, ambition, and the will to dominate may all play some part in indicating competence. Admittedly, these are factors difficult to evidence and to evaluate. Interviews calling for demonstrations of initiative and of capacity to plan, organize, direct, and control may be the best evidence procurable in proving the possession of "traits" usually identified with success in administrative or executive positions.

### *Professional and Technical Positions*

The personal factors best tested orally for positions of a professional or technical nature will vary with the extent and kinds of human relationships involved in the positions to be filled. These factors may be very weighty and yet remain elusive. Great care must be exercised that technicians and professional experts, called upon to participate in the interviewing, do not substitute their judgment of the candidate's technical

or professional knowledge, tested by other qualifying processes, for their evaluation of personal fitness. Too often the maladministration of interviews for this type of employment has brought down wrath upon the entire program of interviewing because technicians or experts who serve as members of examining boards, however competent they may be in their own fields, are amateurs in interviewing. On the basis of one or two questions (which may reflect their pet theories) they may reverse or vitiate the findings of comprehensive written examinations or extensive and careful evaluation of professional or technical training and experience.

### *Law Enforcement Positions*

In classes of work involving law enforcement the use of the interview is again of primary importance because of the complicated personal elements inherent in the ability to obtain confidence, trace irregularities, gather evidence, locate suspects, and bring about convictions. Active policing can be successful only if those engaged in the work bring to the performance of their duties the character and personality which command respect and obedience to the law. Even though we have much to learn about the factors which constitute personal qualifications for any of these positions, we must depend upon the interview, but only with full recognition of its limitations and after employing every known safeguard.

### *Social Welfare Positions*

The interview is necessary in testing for any class of employment concerned with the care and welfare of individuals and families in need of relief, and those who are cared for in public institutions. A rare combination of personal understanding and impersonal conduct, an alert social conscience, and an avid interest in the needs of unfortunates is of primary importance. Here, too, more research into factors of qualification is needed. Experience indicates that the measurement of these factors will probably depend on observing behavior and the subtle revelation of attitudes quite as much as on evaluating the answers to

formal questions. We have not yet attained the same degree of objectivity in testing competence in the social science fields that has been achieved in the physical sciences, probably because data must be dealt with in relative terms. Observation and evaluation of attitudes and reactions are, however, definitely functions of the interview.

### *Labor Classes*

The interview is obviously not well adapted to the selection of the most competent laborers. Many jurisdictions have avoided the problem of testing laborers by such devices as exempting them from the jurisdiction of the personnel agency either in the basic law or by subsequent action of the personnel agency, or by providing for their selection on the basis of priority of application. Many jobs properly classified as laborer require primarily tests of strength, stamina, and skill in the use of the tools of the job; all these attributes are measured by physical examination and demonstration tests, better than by the interview. A short probationary period on the job serves the important purpose of sifting out those who are not inclined to apply themselves faithfully and industriously. Positions classified as "laborer," having requirements that should be tested by the interview, are undoubtedly improperly classified.

### *Clerical Classes*

It is equally difficult to justify the use of the interview for many classes of non-supervisory clerical employees, except when their duties require them to meet the public personally or by telephone, or when the degree of teamwork and cooperation required of employees makes their personal characteristics a matter of primary importance. It is true that in any organization the employees must work together, but the relationships existing in most clerical offices are so varied and so tempered by special conditions that it is not practicable for an interviewer to judge the probable success or failure of an otherwise competent clerk. It is the function of the probationary period or of intelligent supervision to deal with cases in which adjustment becomes a problem. The competitive interview does not

have a proper place in the measuring of candidates for most of the lower level clerical classes.

#### USE OF THE PROBATIONARY PERIOD AS A SUBSTITUTE

It is sometimes argued that the appointing officer will invariably interview each candidate before appointment, and will have ample opportunity to observe the quality of his service during the probationary period.<sup>6</sup> True, he has these opportunities if he takes advantage of them; but there are cogent reasons why this fact does not justify the qualifying agency in disregarding the competitive interview.

In the first place, the appointing officer is entitled and expects to interview certified candidates who are qualified for the position. The personnel agency should therefore take whatever steps are necessary to insure that misfits have been eliminated and that every eligible has been thoroughly examined. The competitive features of an examination are lost if an important part of the testing process is left to the unstandardized, subjective, and inconsistent conclusions of a large number of appointing officers.

Secondly, civil service laws and pressure groups of eligibles tend to protect the interests of those on eligible lists, thus making it difficult for the appointing officer to function freely as part of the testing machinery at the time he makes his appointment. Most civil service laws provide for the certification of three names. Failure of the personnel agency to discharge its own responsibility in eliminating misfits from the eligible lists results in the passing over of those certified who "look as if they may be misfits," and the time soon comes when three such unfortunates stand at the top of an eligible list, either blocking further effective use of that list or imposing upon some department head the obligation of choosing the least of three possible evils.<sup>7</sup>

<sup>6</sup> See the Committee report on the Probationary Period in this series.

<sup>7</sup> Some civil service laws and rules provide that an eligible who has been certified to a particular appointing officer three times and has failed appointment shall not be again certified to that appointing officer except upon his specific request. Such a provision avoids the situation referred to above, but does not lessen the need for adequate initial examining methods.

Finally, probationary dismissals are costly. It is curious that public administrators urge more separations during the probationary period, and at the same time complain about the costliness of discharges. Too often, however, department heads fail to make any use of the probationary period and blame the merit system for the quality of employee they keep. In the face of this known reluctance to separate probationers it may be that a greater emphasis should be placed on the interview to assure the selection of able and congenial employees. But even where probationary tests are fully utilized, economy demands that a department head be furnished, in so far as possible, with persons so qualified that it will not be necessary to discharge them. Any discharge after several months of employment means a loss of far more than the salary of the individual for the period he has been employed. The cost of selecting and training a successor must be added, as well as the effects of an upset organization.

But even though a thorough oral test has been held, probationary testing remains important. No test yet devised can guarantee that a newcomer, however competent, may not become a disappointment under the conditions of new employment. There is no certainty that he will not be poorly adjusted, become discontented, or irritate others. There is still room for a new employee to fail to make good despite the promise he has shown on the basis of written and oral tests and an impressive experience record. Investigations may reveal past shortcomings but the best investigation cannot assure against a future deviation. Both examiners and investigators can make mistakes, and the physical or mental condition of applicants may change. No battery of tests is infallible, and for that very reason the examining process should be as complete as possible.

The probationary period will always be an essential part of the examining process. While the presence of a person's name on an eligible list indicates probability of success so far as the personnel agency can determine it through the means at its command, trial on the job tests factors that earlier steps in the examination may be unable to measure. But the mere

existence of this last test cannot be considered an excuse for avoiding responsibility to complete each of the earlier tests carefully.

#### THE EXAMINER'S RESPONSIBILITY SUMMARIZED

An oral test is the best known method of evaluating certain qualities of vital importance to the performance of work in some classes of positions. The controlled use of the interview as a means of measuring such qualities is good practice.

Criticism concerning the use of the interview comes frequently from disgruntled persons who have not been successful in obtaining places on an employment list and who hasten to ascribe their failure to some cause outside themselves. Criticism, when valid, points toward the need for attention to the way in which the interviews are conducted, rather than to inherent weaknesses in the principle of the interview as such.

The question as to whether or not the interview has proved effective in a given jurisdiction can best be determined by reviewing the practices in that jurisdiction. If the purpose and methods in each test are matters of open record, the examining process is objective and standardized, and the interviewers themselves are competent, the use of the interview can almost always be justified; if these conditions do not obtain, then those in charge of the program are open to attack for improper administration. The device of the competitive interview should not be condemned wholesale because it has occasionally been abused.

Strong arguments against the use of the oral test will be supported by facts in many jurisdictions: the administrators have not sought to evaluate significant factors of fitness; judgment is exercised without evidence; and finally, ratings have been extremely subjective and unreliable, or biased by political, racial, fraternal, or other irrelevant considerations. The apparent conclusion is that the oral test offsets rather than reinforces other parts of the qualifying process. These criticisms, however, are justified only because of the way in which the oral tests have been administered in a given place, and do not reflect

conditions inherent in the testing device. It is a responsibility of every public personnel agency, first, to learn the ways and means by which these faults are to be avoided, and, second, to make so completely known to the world at large the benefits and the methods of a proper interview that further attacks will be minimized.



## *Chapter IV*

### **Principles of Sound Competitive Oral Testing**

THE diversity of practice among various jurisdictions would seem to indicate the need for a discussion of principles, and fundamental limitations upon the use of the interview in the examining process. Such a discussion at this point will add significance to the chapters which follow. Principles and limitations affect—if indeed, they do not control—the integrity of the interview process and the validity of the result. It is the purpose of this chapter to explore the nature of the competitive oral test, the importance of defined objectives, the meaning of factors, the different kinds of evidence which may be developed, possible standards for its analysis and evaluation and for the weighting of factors, and finally to point out again the importance of the record for review.

#### **NATURE OF THE PROCESS**

The oral test, or competitive interview, is a method of developing measurable evidence of the suitability of an individual for a particular position. To be measurable, evidence should be trustworthy and probative of one or more qualities, capacities, traits, or known patterns of behavior or conduct which will affect performance on the job. Methods of weighting the factors of fitness and rating the evidence must be standardized to assure equality in competition.

There are many possible factors of personal fitness, often interrelated, and various kinds of evidence. Some personal factors, separately or in combination, can be evidenced by methods less costly and more objective than the interview. Some factors may be evidenced only through the interview. It is the first duty of examiners to determine which necessary

components of fitness can be evidenced most effectively by means of the interview.

### PRE-SELECTION OF FACTORS

To secure capable employees—the goal of any personnel selection program—requires that the examiner be master of his medium. He must know which tool is called for at each stage of the process, and be able to use that tool properly and effectively. In order to do so, the examiner must first have a knowledge of the job factors involved. He can then plan his approach.

#### *Job Analysis*

A first step in any examining process is to determine the qualifications, patterns of behavior, characteristics and capacities which are *essential*, and those which are *desirable*, for effective performance on a given job. This requires thorough job analysis—not only of the duties and responsibilities pertaining generally to a class of positions, but also of the specific conditions and relationships to be met in each position in the class. Statistical studies of the performance of employees currently or previously employed are often of great aid in such an analysis. The superficial analysis of job requirements is too often a handicap to examiners in preparing and holding an oral test, and a more thorough analysis is called for if the oral process is to have any substantial degree of validity.

#### *Breakdown of Factors*

The second step in any examining process is to analyze the general qualifications, patterns of behavior, characteristics, and capacities—both required and desired—so that they may be allocated to the proper subject of examination. Education, experience, aptitude, skill, knowledge, and personal adaptability are common subjects of examination; each is made up of separate elements.

Each subject may be best evidenced through a particular medium—by means of an assembled or unassembled examina-

tion, through a written test, oral test, or demonstration of skill. Sound procedure requires that the best possible method of evidencing each qualification or capacity be used; but in every case the economic, geographic, and human facilities of the examining agency must be considered in deciding on the appropriate method.

Because of the cost of the interview in time and money, and because of the dangers of subjectivity, only those elements which do not properly lend themselves to other forms of test should ordinarily be included in the competitive oral test. The factors that are reserved for the oral test are usually traits or characteristics of typical behavior, and personal capacities, as distinguished from factors of quantitative experience, skill, and knowledge.

#### *Ratability of Factors*

The next step in preparing the interview is to analyze carefully the traits, characteristics, and capacities reserved for the oral test and break them down into their *smallest component elements*. The items that are sufficiently probative of fitness for the job to justify their inclusion in the interview should next be examined to see whether reliable evidence of them can be procured. In the analysis of factors, as well as in consideration of evidence and rating of candidates, it is important always to remember that fitness is a composite of many elements. Few factors, however relevant, could stand by themselves as criteria of fitness. The minutiae of evidence are ultimately to be weighed as a whole. Traits, actions, and reactions must be considered in combination if behavior patterns and fitness are to be appraised with reasonable validity.

#### *Probative Value*

To ascertain whether a proposed factor of itself or in combination with others has probative value an inquiry should be made as to whether it is a significant component of fitness for the particular position. "Appearance," a frequently used factor, may be an important component of fitness for an in-

formation clerk who meets the public, but not necessarily for a telephone operator. "Voice" and "enunciation" may be essential components of fitness to operate a telephone, but not to serve as an accountant.

Even after ascertaining that a proposed factor is a significant component of fitness for the particular position, it is still necessary to break down that factor into all of its component parts and determine just what quality of appearance or voice is significant and can be evidenced and rated. As to over-all conclusions that a candidate has a "good appearance" or "poor appearance," "pleasing voice" or "harsh voice," honest judges may differ unless they have first agreed as to what combinations of evidence will justify the conclusion. A manner of dress that seems "snappy" or stylish to one may be foppish or in bad taste to another. Thus, examiners must know in advance what elements are significant indications of fitness, and what the standard for evaluation is to be.

Interviews have often sought to evaluate the manner and the judgment of applicants. Both manner and judgment are composites which can be fairly appraised only by an analysis of their component parts. What kind of manner is essential? What manner is suitable or unsuitable on this job? What are the examiners looking for? A "winning smile," or "apparent earnestness," a "firm handshake," or "looking one straight in the eye"? What is the value of each of these bits of evidence? Is each one worth rating alone? Are these items significant in relation to other evidence? Is heartiness, or modesty, or assurance, or vigor sought? If so, which evidences are acceptable as proof of the possession of such factors and how much are they worth in the whole picture as indications of fitness? Are there degrees of value in each factor which can be recognized and rated competitively or comparatively as among candidates? Or is weakness in or absence of the proposed factor a disqualifying element? Or, as is more often the case on analysis, does the proposed factor serve no useful purpose at all?

More and more examiners have come to recognize that many factors heretofore used are only names for abstract qualities

and have little probative value. They have been accepted with little or no consideration of the elements of which they are composed, and of whether or not such elements taken together and with *other* elements are indicative of fitness. For some jobs, certain individual factors may be essential, but for the vast majority of administrative and professional positions observations or conclusions as to appearance, speech, manner, and judgment fail to reveal fitness or unfitness. At most we can say that "good appearance" or "good manner" *may* aid its possessor in performing certain duties, and that weakness in any of these factors *may* be an obstacle. They are not necessary components of fitness except as they may be pieces in a composite pattern; of themselves they usually have small probative force. A candidate who possesses such assets may fail to perform on the job; sometimes he may turn hindrances into advantages. What we wish to rate is evidence of the applicant's characteristic behavior under a variety of conditions, and his demonstrated capacity to accomplish results, i.e., what he has demonstrated he can and will do under various related conditions and circumstances *regardless of potential hindrances or aids* (such as appearance or manner) *which may or may not affect performance*.

### *Can Factors be Evidenced Competently?*

Having determined that some combination of proposed factors is worth testing for—that each is, in fact, a significant component of a whole that indicates fitness—it is still important to consider whether each can be reliably evidenced in the course of an interview. Reliability is relative. Some types of evidence are more reliable than others. Some are so unreliable as to constitute no fair or proper basis for rating in a competitive test. To determine whether or not a factor can be reliably evidenced, examiners must be able to differentiate between the kinds of evidence adduced in an interview, and to understand the attributes of evidence. Unless a proposed factor can be competently evidenced in the interview, it should not be included in the oral process. In the following discussion of evi-

dence some legal terms are employed; but in the conduct of examinations legalism is to be avoided. First, last, and always it should be remembered that the examiner is dealing with human values. The examining process must not become a legalistic end in itself. Nevertheless, in studying this tool of the process, the examiners should make use of the long experience of the legal profession in defining the nature of evidence and its reliability.

### THE NATURE OF EVIDENCE

The principles of the law of evidence stem from simple concepts that have gradually developed in years of experience in the judicial determination of human rights. The primary concept is that evidence offered in the trial of any issue may vary in relevance to the issue and in reliability. Evidence that is irrelevant to the issue, or evidence that is immaterial, unimportant, or so unreliable as to deserve no credence (hearsay and opinion), or which is not the best evidence obtainable is excluded. Such evidence, when offered in a court of law, is not permitted to go to the jury. Upon equally sound principles of justice and logic, untrustworthy evidence should not be used as a basis for competitive rating by examiners.

### *Analogy of the Oral Test*

In an oral test the examiners may be compared to a judge who tries issues without a jury. The examiners must hear the evidence, determine its materiality, its relevance, and its competence to be rated. Having so determined, they must disregard that which is incompetent, irrelevant, and immaterial, and award the rating only on the basis of competent evidence. A combination of evidences may enter into any ultimate determination, and the examiners must make the ultimate determination as to the total worth of admissible evidence. That is, they must award the rating as a judge without a jury awards the verdict. The process is the same and the need to discard untrustworthy evidence is the same. In one respect the examiners' task is the more exacting, for the verdict of the examiner

must not only be right in substance to do justice in each case, but must also in every case stand comparison with the verdict rendered as to every other competitor. That the conditions are difficult is no reason to relax accepted standards of proof.

The purpose of the oral test is to afford the applicant an opportunity to evidence his fitness and ability to do the job, that is, his possession of the personal qualities, characteristics, and capacities to be rated. The burden of proof rests upon him. But the burden of conducting and directing the test, helping the applicant to supply the relevant evidence, rests upon the examining board. The questioning process, like the observing process, should be directed to the development of worth-while evidence—that which is relevant, material, and competent.

### *Relevance*

No small part of the evidence developed during the interview is irrelevant to the ultimate objectives of the test. Applicants and examiners, too, are prone to introduce matter that has no bearing upon qualification, and many times this may be necessary. Irrelevancy may be permitted to clog the record simply because an unswerving requirement on the part of the examiner for legalistic precision in the production of evidence would inject into the process a rigidity and formality which would detract from the result of the interview. Except for the time it consumes, the interpolation of irrelevant or immaterial evidence into the record need cause no concern if the examiner is able to recognize it as such, appreciates its irrelevance, and disregards it in rating. What is relevant it not a matter of first appearance. (Examiners should recall the story of the trial of a milk dealer for watering milk, wherein the court is said to have found that the presence of a trout in a can of milk was relevant to and *prima facie* evidence of the charge of watering!)

In the course of an oral test, evidence of several degrees of value may emerge: that which is of significance; that which is of doubtful value; and that which is of no importance. To illustrate, a candidate for an administrative position may sub-

mit evidence indicating that he excels in sports. The examiner, who may himself be a keen sportsman, must not permit his approbation of the examinee's physical prowess and his belief in the virtue of sports to influence his evaluation of whatever evidence there may be of particular administrative capacities. Similarly, the candidate who shows deficiencies in fields unrelated to fitness for the position to be filled should not suffer in rating because the examiner is personally impatient with these deficiencies.

Evidence is relevant if it has probative value—if it is demonstrative of any one or more of the traits or capacities which in combination will evidence a known pattern of behavior or conduct, or which alone will decisively affect performance of the duties of the position. Certain types of evidence are definitely irrelevant, others are barely relevant, and still other types may have a high degree of relevance. For example, observations of a candidate's dress, peculiarities of speech, apparent nervousness, or evasiveness in answering may all be important per se, for certain positions; they may be wholly irrelevant for other positions except in combination with other evidences of mental condition or behavior patterns which they help to confirm or refute; or they may be of no importance at all in the absence of other elements which would have made them significant. As evidence of demonstrated capacity, for example, a statement by a competitor for an administrative position that his uncle held an important administrative post would be irrelevant, inasmuch as it would have no value by way of showing the applicant's capacity for administrative service. But if the applicant had stated that he himself had handled a variety of problems in an important administrative position, detailed evidence concerning the position, the problems, and his actions and reactions would be highly relevant. If, on the other hand, a candidate's "connections" were deemed a ratable element for a particular job, and if this were announced as a factor (which would be highly undesirable), then evidence about his uncle would be relevant as to that particular factor. Examining boards should, therefore, be constantly alert to keep the objectives or



factors in mind, to limit irrelevant data, to follow up and connect relevant evidence, and in the end have no hesitancy in disregarding all surplus data.

### *Materiality*

Having determined that evidence is relevant, the next duty of the examiner is to consider its materiality. For present purposes the term "material" may be defined as that which, relating to the issue, is of consequence, in contrast to that which, though it relates to the issue, is of no consequence. To be material, the evidence must have that indispensable characteristic of importance; it must not only relate to the issue, but also affect the determination of the issue. It must aid the examiners in evaluating the applicant's possession of some quality at issue, and contribute to an ultimate determination. Obviously, it is not possible to say categorically before the examination is completed that certain observations or reactions or statements, if relevant, are sure to be immaterial. We do not know in advance. But it is of the highest importance that examiners finally distinguish between what is material and what is immaterial to the rating. To attempt to use *all* that has transpired in a test to arrive at some over-all subjective impression, without consciously sifting the record and analyzing the evidence for its materiality and relevance to the factors, is a misconstrual of the competitive method.

Turning again to our imaginary applicant for a position involving administrative responsibilities, let it be assumed that he asserts that not only is his uncle a well-known industrial administrator—an irrelevant statement—but that he, himself, had helped develop policy in his uncle's organization. If, upon questioning, the examiners find that the administrative activity of the applicant in that position was negligible, that, although he was consulted out of deference to his uncle, his advice was regularly disregarded, such evidence would be classified as of little materiality. His reaction to such treatment might be, however, highly material in ascertaining a relevant behavior pattern, but that is a different factor. The examiners must always

be aware of what factor the evidence is related to. If the applicant's statement indicates that his previous employment involved a modicum of responsibility, then it could be classed as slightly more material and the details developed accordingly. The criterion of materiality always is the extent to which the evidence reveals a trait or pattern of conduct or capacity which has been announced as a factor related to performance in the position for which the examination is held.

If examiners do not consciously weigh the materiality and relevance of each item of evidence prior to rating, then the application of a rating standard to the evidence will lack validity. The application of the measuring stick to the evidence should not be based on guesswork or snap judgment. It should be based on sound analysis.

### *Trustworthiness*

To be competent, that is, ratable, evidence must be not only relevant and material to one or more of the factors of fitness but also reliable. Opinions, conclusions, and self-serving declarations are often untrustworthy. What a witness thinks he could do and says he would do under hypothetical conditions (as distinguished from a statement of fact as to what he has done) is not admitted in a court as trustworthy evidence. Statements of this kind are considered untrustworthy in examining also. A candidate may know enough to perceive what should be done in a hypothetical case, and yet have no capacity to do it. He may not dare or care to face the obstacles; he may, in action, take the easiest possible course in preference to the correct one. There is constant danger that examiners will confuse such self-serving declarations of perception as evidence of an ability which may not exist.

However, the answers to hypothetical questions may be competent evidence of *perception* itself—if that is an announced ratable element of judgment—or some other relevant factor. Some examiners believe that evidence of perception can be developed better, or at least more uniformly, by presenting the hypothetical questions in written tests rather than in oral tests.

The advantages of following up, orally, possible alternative answers, and thus delving more fully into the extent of the perception may offset this danger of variations in the difficulty of the problems presented.<sup>1</sup>

A self-serving declaration of intent or will may also be excellent evidence of the candidate's reaction to controlled stimulus, and such a reaction can be a relevant factor in establishing a known pattern of typical behavior or conduct under controlled conditions. Here again the question of what is sought to be proved is of the utmost importance. The objective must be known; the examiner must shape all processes connected with the test toward the objective, and ascertain before rating that the evidence offered is trustworthy proof of the qualifications sought.

While self-serving declarations and reactions to hypotheses may be reliable evidence of conduct patterns in the hands of competent professional examiners, it remains true that, in the hands of examiners not familiar with established behavior patterns, *the most reliable evidence of ability is verifiable testimony as to prior performance—acts and facts—under conditions similar to those obtaining in the job for which he is an applicant.*

The foregoing discussion reveals that there is no real conflict between what have been termed the "legalistic" and the "psychological" approaches to the oral process. Whether examiners are rating factors of demonstrated capacity, or traits, perceptions, and behavior patterns, their ratings have to be based on evidence developed in the interview. A variety of factors and a variety of kinds of evidence can and should play a vital part in the complete test, for personal fitness is a composite of many interdependent elements. Each different approach to the development of evidence requires a knowledge, discipline, and technique of its own. The evidence developed in any approach can be not only reliable but also objective in the truest sense of the term. The real conflict is between the evidentiary and the

<sup>1</sup> See L. J. O'Rourke's development of hypothetical questions in *Oral Tests for Prohibition Agents, Annual Report of the United States Civil Service Commission, 1928*, pp. 35-42.

wholly subjective approach—which is guesswork, if not something worse. Intuitive rating is untrustworthy, for it is admittedly subjective.

### *Objectivity*

Examiners should understand the meaning of the terms “objectivity” and “subjectivity.” The terms are significant in connection with the method of developing evidence, and also with its analysis. Objectivity in examining means that the process is unconditioned by the mind—by self-interest or prejudice, conscious or unconscious; objectivity connotes absence of bias. Accordingly, factual, verifiable evidence is more objective than the opinions or conclusions of a witness that involve his own conception, which is subjective. An examiner’s rating awarded pursuant to a predetermined rating standard is more objective than a rating based on his uncontrolled judgment. That which is objective is more likely to be reliable, but the terms are not synonymous.

### KINDS OF EVIDENCE

In an oral test three major kinds of evidence may be developed:

1. *Observations.* Direct observations may be made by examiners of physical condition, dress, actions, mannerisms, gestures, voice, and the mien of the applicant.

2. *Statements.* Oral statements may be made by the applicant which express fact, knowledge, opinion, conclusion, intention, hearsay, hopes, fears, imagination, fiction, or sometimes downright untruth.

3. *Behavior Patterns.* Behavior patterns, made up of combinations of symptoms, may be revealed by observation of action and reaction and by statements of prior conduct or of intention or belief, which, alone or in combination, spell out to the trained examiner some recognized and statistically established predictable course of conduct. These combinations of action and reaction and of response to stimuli may also reveal to the competent psychiatrist emotional stability or instability, de-

terminable complexes, and other conditions of more or less diagnostic value in predicting fitness.

Of the three types of evidence enumerated, the third deserves special mention. Evidence in the form of a certain behavior pattern is of extraordinary importance because (1) it provides the only known gauge of stability and adjustability for certain kinds of work, (2) it is capable of detailed analysis and statistical validation as fully as any other kind of evidence,<sup>2</sup> and (3), it is commonly misused by amateurs, legalistic fanatics, and psychological fakers. This kind of evidence, therefore, however valuable in competent hands, should not be rated unless experts are available, who are also trained and willing to subject themselves to the discipline of standardization required for competitive civil service testing.

Each of the three kinds of evidence mentioned will come before the interviewers in the course of an oral test, and must be anticipated in the preliminary analysis of factors to determine how the factors proposed can best be evidenced and rated. Only after such a pre-analysis of the nature of evidence to be developed is it possible to make an intelligent pre-selection of factors or issues to be rated in a particular test.

The problems of developing evidence during the interview, of analyzing it for rating, and finally of applying the rating standard, are the ultimate problems of oral testing; but, before considering further the principles underlying each of these, references should be made to the problem of assigning proper weights in oral tests to the several factors which, alone and in combination, are to be the determinants of personal fitness.

#### WEIGHTING FACTORS

To weight factors is not easy. Isolated factors, although probative of fitness, and established by evidence (however "necessary," as distinguished from "desirable," they may seem), are usually inconclusive in determining the candidate's fitness. It

<sup>2</sup>It should be noted, however, that this type of evidence is unfortunately often abused, for those who are capable of interpreting it properly are seldom also trained in the discipline of competitive testing.

is the composite picture that is revealing, and each factor is only a component of that picture—a component of variable significance under variable conditions.

The value of a factor may be great or small, positive or negative, depending upon its consistency with other data about the individual. Moreover, the importance or weight of any element is likely to vary with the degree to which it is manifest; many elements, which have little significance within the usual limits of variance, may, if wholly absent, or present in extreme degree, be overwhelmingly important. For example, a total lack of ambition might predicate indolence; excessive ambition, unscrupulousness; whereas a normal degree of ambition would be unimportant.

Nevertheless, in a competitive test weights must be established at some point. At the very least, a standard method for assigning weights, as well as ratings, must be developed and followed. That is, the weights given to each factor when comparably evidenced, or to comparable combinations of factors when comparably evidenced, should be the same, and there should be a standardized method by which such weights are determined and which is itself statistically defensible and demonstrable.

### *Fixed Weights*

The fixing of an invariable weight for each factor in advance of the test is easy and certainly standardizes the weighting demonstrably. Too often, however, it is statistically indefensible.<sup>3</sup> How, then, should examiners seek to solve this problem? First, the essentials involved should be analyzed on the premise that factors are elements of fitness, some of which in any test will be of greater significance or weight than others.

The degree of significance will be affected by a variety of influences. The degree to which a factor is possessed, as revealed by the evidence, would be one such influence. It has been

<sup>3</sup>Dr. Arthur W. Kornhauser of Chicago, in a critique presented to the committee preparing this report, pointed out that fixed weights generally adopted by civil service technicians in advance of the interview are "psychologically and statistically unjustifiable."

pointed out also that a factor may be very important—positively or negatively—if possessed in an unusual degree, but has no significance if possessed within a normal range. It may also have more or less significance if *other* factors are simultaneously evidenced in high or low degree.

Some jurisdictions, recognizing this problem, assign the usual fixed weights to the several factors considered important, and then provide a final factor, having an overwhelming weight, to represent the over-all or composite conclusion of the examiners. The examiners, considering all the evidence of all the factors and weighing the degree of possession of each, and the relationship of that degree of possession to the degree of possession of every other factor, then take into account all the variables mentioned to rate the final “over-all factor” as they see fit. Thus they influence or control the final result. While this is a recognition of this problem, it is an avoidance of standardized procedure for weighting and rating. It is, in effect, an assertion that unstandardized “judgment” is better than the formalism of inflexible standardization. The initial weights in such a case have little meaning, for impression and intuition are allowed to exert an over-all control. This avoidance does not offer a solution of the problem.

Other jurisdictions, by assigning fixed weights in advance, decline to recognize the problem at all. Generally speaking, however, too little time and thought is given to the difficult problem of assigning valid weights to factors in an oral test. This despite the fact that the oral test itself may be given substantial weight in the whole examination.

### *Variable Weights*

It requires a great deal of time in advance of each examination to foresee the variables which should be taken into consideration in determining the weights to be assigned, not only to each test, but within each test. But foresight is possible and a standardized method should be attained. If the psychologists are correct, if their conclusions based on variables can be right more often than wrong, some tangible basis for that pre-

dictability must exist. Indeed, in some cases the psychologists have statistically demonstrated the validity of certain of their conclusions as to predictable behavior based on the variable weighting of indicators (factors) possessed in varying degrees and in various combinations under comparable conditions. L. L. Thurstone, at the University of Chicago, is doing interesting work on identification of evidences predictive of abilities, and L. J. O'Rourke, of the United States Civil Service Commission, is studying variable factors of executive ability.

It is recognized that the system of variable weights here discussed is a departure from common practice. It appears, however, to be the only sound approach to the weighting problem, for, in arriving at the ultimate conclusion of fitness, the weight given to the factors is obviously just as important as the rating given to the evidence. The examiner's task is to shut the door to unstandardized over-all judgment without reducing the interview to a legalistic pretense by adopting fixed weights in the name of "objectivity."

The fact that unstandardized judgment may more often be right than wrong in eliminating some obvious misfits, and even increase the probability of selecting the fittest, cannot justify the tolerance of guesswork in open competition. In so far as statistically justifiable results can be obtained by giving different weights to factors possessed in varying degrees and in various combinations under comparable conditions, it is desirable for the personnel agency to develop standards for such differentiation of weights for the factors in each test. At the same time, it must be remembered that the only method by which the assignment of variable weights can be validly standardized is on the basis of statistically demonstrated experience. Such experience does exist with reference to established patterns of predictable behavior based on combinations of traits, characteristics, and capacities in some fields of work. But such experience definitely does not exist at present in many fields of work with reference to many other variations and combinations of traits, characteristics, and capacities.

While it is not yet possible to make extensive use in most



tests of the theory of variable weights here presented, it is time that research efforts be directed to the problem. For most positions, fixed weights will probably continue to be used until scientific study has uncovered adequate data to establish sound standards of variability.

### METHODS OF DEVELOPING EVIDENCE

At an earlier point in this chapter the three types of evidence were described. It is to be expected that each type will be adduced to varying degrees in a given test situation. But unless positive steps are taken to assure uniformity of coverage throughout the particular test, a condition may easily arise wherein the examiners focus their efforts at bringing out one type of evidence at the expense of the rest. This warped coverage may characterize the test as a whole, or it may crop up intermittently as particular applicants come before the examiners. Thus it is desirable for those planning the oral test to determine in advance the extent to which the requirements for the particular position call for the adduction of each of the three types of evidence. Having made this determination, the planners can then develop the proper approach. The following discussion will dwell on methods whereby each type of evidence may be brought out. The use of these methods will be helpful in maintaining the desired degree of uniformity throughout the test.

#### *Observing the Applicant and his Controlled Reaction*

Examiners conducting interviews have before them certain evidence which is presented with the applicant himself—the clothes he wears, the emphasis with which he speaks, his fluency of expression, his physical actions and reactions—all observable matters of fact. When such data, alone or in combination with other known facts about the individual, are relevant to the announced factors of qualification, notes (taken manually or dictated) should record the factual observations.

After the primary evidence, in the form of the applicant, has been withdrawn these notes will serve the purpose of preserving a *factual* record of observations of the candidate and the grounds

for the impression made by the candidate on the examiners. For example, a dirty face, unkempt hair, and baggy trousers are observable facts; an "untidy appearance," on the other hand, is an impression. Recent trends in oral testing technique indicate that care should be devoted to the development of this "observation record," as distinguished from more commonly recorded "opinions" or "conclusions" of examiners which often fail to reflect the basis for the examiners' impressions.

To recognize the difference between the statement of factual observations which support rating conclusions, and the statement of the examiners' interpretation or conclusion from evidence that has come and gone, is of first importance in the task of recognizing competent evidence for rating. The recording of direct observations is the simplest of all methods of developing evidence, and yet it is commonly bungled by interviewers who do not distinguish between conclusions and the facts which support them.<sup>4</sup> So long as examiners substitute their opinions for observations in developing a record to support the rating, this highly valuable implement of measurement will continue to be a mistrusted stepchild in the family of tests.

### *Eliciting Facts About Past Demonstration of Capacities*

The second kind of evidence, verifiable statements of fact as to what the candidate has done, evidencing possession of particular capacities demonstrated in the course of prior experience, represents a comparatively new technique in oral testing. Its successful use by the United States Civil Service Commission has already established its importance as one tool, among others, to disclose possession of personal qualities that cannot be evidenced effectively by the written test.

In developing this kind of evidence, the first function of the examiners is to aid the candidate to place upon the record complete and truthful details as to what he has done. This record demonstrates possession of the capacity or qualification at issue. What the candidate might do under given circumstances (hypo-

<sup>4</sup>A record of observations is particularly important also when examiners are drawing forth observable reactions to stimuli, both physical and mental, which predicate known patterns of typical conduct and behavior—a topic to be discussed later.

thetical cases), what he thinks he ought to have done, what he would like to have done, what he thinks he can do, may constitute evidences of other factors (perception or ability to analyze), but these cannot be considered competent evidence of the possession of a particular capacity. *The best evidence of possession of a capacity is evidence that the candidate has demonstrated possession of it in past action.* The details of such demonstration in action under relevant circumstances constitute the evidence. Examiners should seek to place upon the record the circumstances surrounding each relevant incident described as a setting for the probative facts—what the candidate did under the circumstances.

What did he do which shows initiative or perseverance, if these are factors? What did he do to obtain cooperation, and what happened? What considerations did he weigh before making some decision, and what decision did he make? How long did he vacillate? What was the effect of the decision made? What did he do in preparing a plan? What points did he consider? What did he do to carry through a program once planned, or to adapt the program to changing conditions? What obstacles did he meet? What did he do under the circumstances? Facts, facts, and again facts must be sought out as evidence of demonstration of the capacities selected as factors.

It is not necessary that every example cited of past demonstration of capacity be an important one. Rather the candidates should be impressed at the outset with the fact that minor or unimportant episodes often significantly demonstrate the possession of the required capacity. As evidence of demonstration of the possession of ingenuity, it may be as significant under certain circumstances for a stranded motorist to persuade a taxi driver to lend him a dollar to buy gasoline as to have persuaded the President of the United States to issue a proclamation. Facts and circumstances, not news value, constitute the evidence of personal capacity.

In seeking to aid applicants to place the factual account of their experience upon the record, the examiner may address the same introductory questions to each competitor. Each question is designed to unlock the gates of reticence and memory. Some

candidates offer their experience reluctantly. The examiner must take into account any strain the applicant may be under and should encourage confidence to the end of developing the facts which, under other circumstances, the applicant might readily submit. At the same time, the examiner must be able to direct questions to the less modest applicant in such a manner as to make clear in the record which statements are exaggerated. It is the examiner's function to adapt his methods to the applicant's facility in presentation and to secure a record of factual data which will not reflect variations in the applicants' ability to present evidence of experience.

Another way to assist the candidates in presenting their relevant experiences is to notify them, well in advance of the examination, not only of what factors or qualifications are to be the subject of examination, but also that they are to be asked for examples, taken from their past experience, to demonstrate the possession of these qualifications.

When the candidate's statements are placed upon the record as direct evidence, the examiner is free to follow up the statements offered and to enlarge upon relevant details as to time, place, and circumstances. Cross-questioning is permissible. This does not mean that the candidate is to be heckled; he must simply be led to give detailed descriptions of the situations presented, with time, place, and names connected. By using this method a good examiner is quickly able to disclose misstatements or exaggerations. Having developed the details of place and circumstance, and the names of others familiar with each incident, further verification, if needed, can be made by field investigation. By this process factual and verifiable evidence can be developed, and the evidence so obtained can be recorded. This complete record then contains evidence demonstrating possession or lack of the requisite factors and is a reviewable basis for the application of the rating standards.

### *Eliciting Evidence of Lack of Capacities*

One objection to oral tests has been that, while they often adduce evidence of affirmative qualifications, they seldom ad-

duce reliable evidence of negative characteristics. If no evidence of demonstration of a required capacity can be produced, this fact constitutes negative evidence. The process here described seeks to prove also lack of capacity; if followed through, it produced evidence of the *lack of possession* of required qualifications.

In other words, the process goes further in revealing lack of qualifications than the mere absence of affirmative evidence. For example, in an examination for personnel director for the federal service, the first factor to be evidenced in the oral test was possession of initiative, and the candidates were asked to give examples of the display of initiative in their past experience. They were encouraged to present profuse examples of ideas, plans, and procedures which they had initiated in order to evidence fully the extent of their initiative. A subsequent factor was the capacity to carry through a program to a successful conclusion. Here the examiners were able to refer again and again to many ideas, plans, and procedures which the candidates had stated they initiated, and to ascertain whether they had carried them into effect or had simply thought them up and then allowed them to lapse. When any example of initiative had not been carried through, the reasons and circumstances for the failure to carry through were placed upon the record; this process introduced the candidate to further questions of qualification, such as ability to obtain cooperation and to adapt programs to changing conditions, and so on. The applicant having made original assertions of considerable capacity for initiative, and having provided details as to time and place, the door was wide open; it became relatively simple to disclose failure to carry through, which may have been due to lack of ability, lack of will, lack of concern, lack of daring, or fear of consequence, or, in some cases, due to conditions entirely beyond the control of the applicant.

### *Cautions*

In developing evidence of the lack of certain capacities, it is important that the interview be constantly directed toward the

demonstration of the particular capacities at issue. The questions asked must be pointedly relevant, limited to facts and circumstances surrounding the appropriate examples offered; every excursion of the testimony away from the issue should be led back to the relevant channel. This task of the examiner is of the utmost significance; the time allotted to disclosure of ratable evidence is precious and should not be wasted on irrelevant matters.

It is necessary, however, to include in the record the surrounding circumstances—time, place, and conditions—of each relevant episode. The examiners must make the record complete and follow each incident through in sufficient detail for verification. In most cases there is no need for field investigation to complete verification; the examination is sufficiently exacting. Should investigation be needed, examiners should make a note on the rating chart of any items of evidence which should be further investigated.

Some candidates may never have demonstrated possession of certain capacities for the reason that they have never had an opportunity to demonstrate them. This is not a conclusive objection. It is a limitation inherent also in quantitative rating of experience. In examining for positions wherein prior experience is a prerequisite, applicants who have not had such experience have long been passed over, even though they may have necessary ability. Economy of operation has made that practice necessary, but there is reason in the practice also. As between a candidate who may be potentially capable but has never had occasion to demonstrate the capacity, and one who has demonstrated his possession of the capacity successfully, the latter is the safer choice. This assumes, of course that there is no evidence to show that the applicant has retrogressed markedly since the time at which he demonstrated his capacity. Thus, recency of the demonstration may be a relevant consideration in rating.

During the conduct of the oral test, the examiners must constantly be on the alert for false or exaggerated statements. If the exaggerations stay within reasonable bounds no action in rating should be taken beyond an honest appraisal of the actual facts,

stripped of their embellishments. However, it is the duty of the examining board, within the limits of its facilities, to ferret out what seem to be misstatements. Experience has shown that when questions, designed to develop the finer details of dates, names, and places, are directed to actual experience, even the most ingeniously preconceived fabrication can usually be discovered. Only the most resilient imagination can withstand detailed questions concerning something that existed not at all or only in part. If doubt still exists concerning the applicant's veracity at the end of the interview, specific statements may be adequately checked by investigation.

Before turning to a consideration of the third type of evidence, it is important to mention again the fundamental need for the examiners to control the flow of the evidence. As evidence relevant to any factor is developed, there will inevitably crop up items that are not relevant to that factor, but relevant to other factors or to over-all fitness. Accordingly, provision should be made to note, for further development at another moment and without causing deviation from the objective at hand, those facts or leads relating to objectives not immediately under consideration. It is a good practice to attach to the rating form a separate sheet for noting evidence as to each factor to be considered. On these sheets should be noted for further reference any possible evidence as to a factor as it crops up. Such a sheet is particularly useful in marshaling items of evidence related to particular patterns of conduct or behavior as they are noted in the course of the interview. Confusion results if the examiners try to pursue and develop the evidence relating to a number of factors at random and without order or direction. Each factor should be the subject of exploration at a given point in the interview, and incidental or collateral probing of other leads should be avoided at that point.

Although the applicant presents and is finally rated on an integrated combination of qualities, systemized examining separates these qualities for purposes of obtaining and weighing the evidence. The separation of factors is the only sound basis for objective rating whether the rating is finally awarded for

each factor separately, or for many significant but disparate items which together indicate a pattern of behavior or conduct.

It is because a great variety of evidence will come before interviewers in the course of an oral test that this conscious marshaling of evidence is important. It is the duty of the interviewers to determine what bits of evidence are relevant, material, and competent as to each factor or pattern, to use this evidence for evaluating that factor or pattern exclusively, and to discard irrelevant matter from consideration. Evidence developed at any point may be relevant to several factors and to the composite picture. But for every rated conclusion, evidence should exist and be earmarked and consciously appraised. In this way, examiners can proceed objectively in applying a rating standard, and also in considering the interdependence of all components.

The separation of objectives in developing evidence is particularly stressed because there is a definite tendency in oral testing to arrive at an over-all conclusion from first impressions, and to overlook important evidences concerning particular factors which affect and may or may not justify the conclusion.

Over-all fitness for any job is not extraneous to the characteristics, capacities, and reactions that are its components. It is a composite of many attributes. Some of these attributes or factors may be essential in themselves, for a particular job; more will be dependent upon the existence of other factors. The point is that fitness is revealed by the evidence developed and it is only by the separation and analysis of evidence that the judgment of fitness can be substantiated. The over-all conclusion as to fitness is ultimately arrived at only by evaluating the evidences of many factors. Since the individual is a composite product of many variables, it is necessary before multiplying to isolate the multiples.

### *Recognizing Symptoms of Predictable Behavior Patterns*

The third kind of evidence consists of items which combine to indicate typical behavior or conduct patterns. It is in isolating symptoms of these predictable patterns that examiners face the



greatest difficulty in preserving objectivity. It is easy to throw aside the burden of isolating the evidences which in combination justify a conclusion. It is easy to derive from the whole interview an unsubstantiated impression which may or may not have competitive validity. Yet, if the impression or conclusion has validity it is because there has been developed evidence which, alone or in combination, supports the conclusion on the basis of statistically demonstrable experience. Otherwise, the product is pure guesswork. Excessive concentration on legalism or formalism is likewise to be avoided, just as a traveler avoids hampering his over-all vision by prolonged study of each blaze along the trail. But it is necessary to follow each item of relevant evidence along the trail to valid conclusions; otherwise the examiners may become lost in a jungle of subjective impression.

In this connection, Dr. Kornhauser says of behavior patterns:

It is the composite picture which is revealing. Each fragment of evidence . . . fits into the gradually evolving size-up of the person. Its meaning is great or small, positive or negative, depending upon its consistency with other data about the individual.

In seeking this third kind of evidence, the examiner must remember that the individual items may be insignificant in themselves. Clearly they cannot be separately weighted without regard to one another. Symptoms that make it possible to predict a typical pattern of behavior or conduct are infinitely varied and numerous, and there is an enormous range of functionally equivalent indicators. It is not feasible to compare candidates simply by lining up parallel items of evidence; but candidates may be justly compared on the basis of conclusions inferred from combinations of varied fragments of evidence, at least in so far as the predictive significance of such combinations has been statistically demonstrated.

To draw forth the many significant fragments of evidence to be pieced together, it is necessary to give candidates the fullest opportunity to reveal themselves not through a discussion of their experience alone, but of their hopes, fears, and interests as well. Competent examiners, familiar with statistically valid

conclusions from various combinations of trait-symptoms, will be able to recognize, distinguish, and record relevant symptoms and to arrive at a valid conclusion.<sup>5</sup> If examiners are not so qualified by training and study, they should leave this third kind of evidence alone, or frankly admit that theirs is guesswork. Unable to demonstrate on what grounds their conclusions may be "right" or "accurate," they can only assert that their intention is honest and urge that the exercise of uncontrolled judgment increases the "probability" of selecting fit candidates. Practice in probabilities should be limited today to the selection of applicants noncompetitively. For it is only when relevant evidence which supports the conclusion has been developed and recorded that competitive rating is justified.

### BASES OF RATING

Having developed the evidence related to selected factors of fitness in the case of each candidate for a particular job, the problem of the examiners becomes one of evaluating or rating the evidence to determine the relative qualification of the different candidates. The process of rating should be standardized by standardizing the method (1) of marshaling and analyzing the evidence of the several factors, (2) of assigning weights to the factors as evidenced, alone or in combination, and (3) of defining and applying a scoring system. This succession of steps in the examining process may appear complicated and time-consuming. But it has been observed that, once standardization of method for each of these purposes is adopted in any examination, trained examiners can follow the method without appreciable variation or delay.

### *Standards for the Analysis of Evidence*

Experience indicates that one of the most difficult problems in any technique of oral testing is the proper marshaling and

<sup>5</sup> For a discussion of an interesting experiment, wherein candidates for police positions are subjected to sudden stress situations in the course of an interview and their reactions are observed and rated, see G. L. Freeman, G. E. Mason, E. T. Katsoff, and J. H. Pathman, "The Stress Interview," *Journal of Abnormal and Social Psychology*, October 1942, pp. 427-47.

analyzing of facts and symptoms developed during the interview. Even the most scientifically administered oral test can fail in attaining its objective, if final ratings are applied to irrelevant evidence, or applied in the case of different candidates to evidence having varying reliability or materiality. The competitive principle should pervade the whole selection process, of course, but its most precise application is essential in the interview because of the inherent danger of subjectivity. This danger is greatest at the time the evidence adduced in the interview is marshaled and rated, that is, when the interpretation is made and expressed in a score. A standardized method of rating must be used constantly and used by an examining board that possesses an intimate knowledge of what it is rating.

We need, then, a standardized method of assembling and earmarking for evaluation the evidence that has been offered. Definitions describing precisely the qualities the examiners are seeking may be agreed upon, but unless all the evidence illustrating possession of those qualities, whether separately or in relation to one another, is recorded and consciously considered in assigning ratings, the results of those ratings will be of doubtful value. This marshaling or assembling of all the evidence connected with each ratable element is the first step in rating.

It is maintained in this work that a disciplined procedure for classifying the assembled evidence according to relevance, reliability, and materiality has distinct advantages. Fixed meanings for these terms permit the uniform consideration of the value of evidence.

If the evidence has no probative value, or if it is wholly unreliable or immaterial, it should not be rated. If, however, it has some probative value in terms of a factor of qualification or a pattern of predictable conduct, it should be considered. The degree of its value may vary from case to case. So, also, in determining materiality and reliability, the evidence adduced may vary from that which is slightly material or fairly reliable to that which is significantly important and that which is trustworthy beyond doubt. If a single item of evidence, directly relevant to a particular factor, and wholly trustworthy, is also

highly material, or if there are sufficient moderately relevant, trustworthy, and material items offered, the evidence on that factor would earn the highest rating for that item, i.e., the evidence supports the conclusion that the candidate possesses the quality sought in high degree. Thus, if there is offered sufficient evidence which is material, relevant, and trustworthy, the applicant has demonstrated his possession of the quality under consideration and the evidence warrants awarding the applicant the full credit permitted for the factor alone or in combination. Contrariwise, it is obvious that if all of the evidence developed as to that factor is irrelevant, immaterial, and untrustworthy, the applicant has failed to demonstrate possession of the quality and should receive no credit for it.

A method of analysis of all evidence greatly aids the examiners to reach a sound and supportable conclusion. It may not be necessary to provide a separate analysis scale for each candidate, as the United States Civil Service Commission has done by way of experimentation in some important tests. The danger of excessive formalism is always present. For those interested in the possibility of developing analysis records by formulae, the following note, prepared by James C. O'Brien, will indicate an experimental approach to the problem:

By the maintenance of check marks on an analysis scale for relevance, materiality, and trustworthiness during the interview, specific data is at hand for application of a rating scale to the evidence. This analysis scale should provide space for the notation of each item of evidence offered by the applicant, and for checking as to each item the degree of materiality, relevance, and reliability. Suppose, for simplification, the symbol "R" denotes relevancy, "M" signifies materiality, and "T" represents trustworthiness. By the use of these symbols each example or major item of evidence could be classified or graded by degrees under each basic symbol. Thus, each item of evidence could be classified as "R<sub>1</sub>," "R<sub>2</sub>," "R<sub>3</sub>," or "R<sub>4</sub>"; "M<sub>1</sub>," "M<sub>2</sub>," "M<sub>3</sub>," or "M<sub>4</sub>"; and "T<sub>1</sub>," "T<sub>2</sub>," "T<sub>3</sub>," or "T<sub>4</sub>." If the example offered is slightly relevant, highly material, but only moderately reliable, the checked scale to which the rating would be applied, might read as follows: "R<sub>1</sub>M<sub>4</sub>T<sub>2</sub>."

I shall cite, as the simplest possible example, the analysis of evidence developed in an oral test relating only to administrative ex-

perience. Let it be assumed that the announcement proposed to rate the quality of the best 3 years of experience of candidates in *duties comparable to those of the job offered*, and that a candidate submitted an instance in his employment record where he had actually performed, for the requisite duration, what he called "comparable duties," as an office manager in a small manufacturing establishment where a considerable number of production workers came under his supervision. The examiners would elicit by careful questioning the nature of his duties and responsibilities on that job. It may develop that the duties and responsibility of the job were similar to those of the position to be filled, and that they involved administration and supervision to a degree sufficient to meet the requirement of relevancy. Let us further assume that on the face of it, the applicant's supervision of a large number of employees represented a broad supervisory responsibility as such. But because a foreman actually handled all details of supervision on the job and he did not, his supervisory experience was less significant in proof of qualification for this job than it might be. Since the job *required* similar personal qualities and responsibility, this evidence would be classed as highly relevant with the symbol " $R_4$ " attached thereto. Since a fortuitous fact had prevented the use of his potential opportunity for experience in supervision, the evidence would be classed as " $M_3$ ." Since the factual recital was complete in detail, without contradictions, agreed with the record of employment and compensation, and the stated facts were confirmed by vouchers, it would be rated as " $T_4$ ." As a consequence, the whole example would be characterized on the analysis sheet as " $R_4M_3T_4$ ," the value of which, in terms of score for *this experience*, would be fixed from the *rating scale*. Of course, judgment on the part of the examiners is still needed in applying the formula, but the subjective element in scoring is reduced to the minimum by the use of this analysis. The significance of the "valences" assigned to the various symbols may be determined in advance of the oral test for use if necessary by different rating boards. For example, it may be decided in advance of the examination that if the letter "R" is not found in a sufficiently high degree (predetermined) in a sufficient number of the formulae attached to the applicant's record of evidence related to a given factor passing credit is not justified. The applicant has failed to prove his case through the presentation of relevant evidence. It may also be decided that if, in any case, there is a consistently low "valence" of the letter "T" the matter should be scheduled for a subsequent personal inquiry or investigation to verify the statements made during the oral interview. By discussions and agreement in advance of oral tests, definite bases for applying the rating scale to different

values of evidence received may be determined; they would constitute a precise standard for an appellate body or anyone else who may be interested in subsequently considering the examining board's work.

It goes without saying that the foregoing discussion of reducible formulae is offered only as a mechanical approach to the establishment of standards of analysis of evidence. Mechanics can be overdone. But it is important that examiners emphasize the necessity for an accurate and conscious typing of the wide variety of evidence received in all oral tests to accomplish the transferral of rating from the quicksands of fleeting impression to the solid ground of reviewable fact.

### *Standards for Applying Variable Weights to Factors*

The practice of fixing in advance of a test the weight to be given the rating of each factor in the final result has become common. The predetermination of a fixed weight for each factor unquestionably establishes weight standardization. But the need for a system of variable weights, dependent in part upon the evidence developed, becomes apparent in view of the consideration that some factors are of importance in the final determination of fitness only in relation to others, that some are of little importance when they are possessed in a normal degree, but of great importance, positively or negatively, when possessed in an extreme degree or not possessed at all.

Unless some standard method can be worked out in advance of each test, to control the fixing of the weight of the factors, to be applied after the evidence has been developed, variable weighting will open the door wider than ever before to conscious or unconscious favoritism. Without standardized controls there would be little uniformity by the time the hundredth candidate was reached. Weights finally assigned in each case would depend upon the examiner's impression of the candidate as a whole, without regard to comparable elements and combinations, or to the specific evidences in the case. However, there is still insufficient statistical demonstration of the value of different combinations of factors possessed in different degrees to arrive at a standard for fixing variable weights.

A five-point scale is used to indicate the degree of possession

of the several factors. It is possible to assign a high, an intermediate, or a low weight to a factor, depending upon the degree of its possession. It is also possible to assign a variable weight to one factor so that it will have a higher or lower weight, depending upon the degree of possession of other factors. In short, a basis for assigning variable weights can be stated in advance of the test which will take into consideration significant variances in data disclosed during the interview. By stating in advance of the test the different weights which may be assigned to each factor, under different stated conditions and combinations, a standard can be established.

A standard should not be established, of course, unless it is valid. The fact that fixed weights may not have validity is not an excuse for adopting a more complicated and equally invalid procedure. When the variations and their treatment can be shown to have statistical validity in experience, the effort is worth while. At present such a demonstration, while possible for some types of work, remains the exception rather than the rule. Before standards can be propounded for most classes of positions research must be done; therefore, extreme caution should be used in experimentation with variable weights. Their employment should be limited to the few recognized psychological and behavioristic combinations which have been statistically demonstrated. However, the greatest hope of establishing sound competitive evaluation of the complex elements of personal fitness appears to lie in this direction.

### *Standards of Rating*

In applying any rating scale to evidences of a candidate's fitness, whether the evidence has been developed in reply to written questions or by oral statements or observations, examiners have to use their training and judgment in a majority of cases. It is apparent that standardization of the method of analysis of evidence and of the values to be assigned to different factors can reduce the probability of error; but these first two methods of standardization do not assure uniformity in the method of applying a rating scale. It remains necessary to standardize so far

as possible the method of applying ratings so that variances in the final use of judgment as between candidates, and types and combinations of evidence, will also be reduced.

Uncontrolled impressions and conclusions concerning elements of personal fitness need not be tolerated, because relevant and reliable evidences can be developed in an interview. Likewise, uncontrolled evaluation of evidence need not be tolerated. The day of subjective judgment as to the value of subjective impressions formed by oral examiners is definitely past. Not only is double judgment tabooed, it is possible now to make uniform the method of evaluating objective evidence for each candidate.

The first step toward the standardization of rating method is to define precisely the factors themselves and the ratable evidences of their possession. The probable variations in the interpretation of observable factors such as "appearance" have already been discussed. The same is true of capacities. Consider the factor "initiative" (a single component of administrative ability). Initiative may mean one thing to one examiner and something else to another. One may think of initiative in terms of the quantity of new ideas conceived; one may think in terms of the "push" or "drive" of the candidate; one may think of initiative in terms of the quality of usefulness of ideas conceived. Without definition, each examiner may look for a different kind of evidence, or evidence of something different than what the others seek to rate.

Some jurisdictions define each factor, the kind of evidence to be rated, and the bases of the ratings (when more than one basis is significant) on the face of the rating form to be used for each factor. For example, the United States Civil Service Commission, in an oral test for personnel director, defined the factor "initiative" on the rating form, and provided two rating scales, one quantitative and one qualitative. Each scale had a different weight. The definition follows:

Initiative—Evidence desired: Examples of demonstration of resourcefulness in devising ways to meet particular needs of the individual, his work, or activities; examples of original proposals or



changes initiated by him; and examples of ingenuity actually conceived.

Below this appeared the two scales with the following notation:

The rating of this factor is *quantitative* on the first scale for profusion of ideas initiated, it is *qualitative* on the second scale for the constructive quality of conceptions, the ingenuity shown.

In this way, the approach of examiners is standardized so far as their first objectives are concerned. The definitions and separate scales assure consideration of each item of evidence in relation to defined objectives. The usual method of rating "initiative" confuses the questions of quantity and quality of initiative evidenced. Some individuals may be extremely profuse in new ideas but the ideas may be worth very little. Yet, as the rating scale itself indicates, an individual who is thinking in terms of original ideas may be worth more for a certain position than the individual who never has new ideas. Likewise, ingenuity and the constructive quality of a single idea developed may have an important bearing on qualification for a particular position regardless of the lack of "profusion" of original thinking.

### *The Rating Scale*

The second step toward the standardization of rating method is to define the scoring range or scale to be used. There are various possible methods of scoring, such as the awarding of credits on a scale of 100; or the assigning of candidates to various categories under each factor (such as the five-point scale: "excellent," "above average," "average," "weak," and "poor"), with a numerical value pertaining to each category; or the awarding of increments or decrements for each factor to the preliminary rating on experience. The most that can be done to standardize the method of award under any one of these systems is to discuss with and write out for the examiners the definition of the several degrees of possession of the factor expected for each grouping on the scale. Specific examples typical of each value in the

scale are here desirable. Analysis-formulae are sometimes helpful.

The problem of developing an adequate rating scale presents the ultimate problem of reducing subjective judgment to a minimum. Having established definitions and sample analyses of various types, items, and degrees of evidence which may develop to establish factors, examiners will have gone as far as the present state of our knowledge permits in securing objectivity and uniformity.

It must be recognized that the value of each possible item of evidence, either alone or in relation to other items, can never be reduced by formula to an exactness which can be mathematically stated on a rating scale in advance. Exactness of that kind has not yet been demanded, and it is no more necessary in interview rating than it is in rating either performance, or free-answer written tests. In view of these considerations, then, the maximum of desirable standardization will be obtained under the following conditions: (1) Relevant evidence as to each factor is adduced and preserved on the record. (2) There is adequate definition of the rating objective. (3) There is adequate analysis of evidence. (4) There is adequate analysis of the basis for the weights to be given to factors evidenced in different degrees, and in different combinations. (5) There is adequate definition of the scoring method. These safeguards assure, so far as it is now possible to do so, a common standard for the evaluation of competent evidence and minimize the probability of subjective error.

#### ESTABLISHMENT OF THE RECORD

The final essential of a sound competitive interview process is the establishment of a full record of the procedure used: a statement of the reasons for using each factor selected and its relevance to performance required on the job, the definitions agreed upon, the method and basis for assigning the weights to the factors, the common questions or directives addressed to each candidate relevant to each factor, the evidences developed by both observation and statement, the method or standard of scoring adopted with definitions, and, last, the procedure used in

evaluation (discussion of ratings, averaging, and the manner in which the rating standard is applied to the evidence). All of these points should appear clearly and separately in the record.

The examination announcement is, of course, part of the record, but the announcement is often inadequate in respect to the definition and weighting of the factors and the standards set for analysis and evaluation. Further data—all of the definitions and determinations of the examiners—should be duly recorded and preserved in the examination file. Various methods of preserving a record are discussed in the section of this report that deals with the administration of oral tests. This more detailed discussion serves further to underscore the necessity for compiling and retaining a full record of the whole oral test process—not solely for the sake of fairness to competitors, but in order to make it possible to review the test.

## *Chapter V*

### **Administrative Aspects**

**I**N A central personnel agency the oral test or competitive interview is part of the total examining process. It is common practice to place responsibility for such interviews in the examining division subject to the direction of the chief examiner. This practice is based on the sound principle that all parts of the examination should be under central direction and each part of the testing process carefully related to every other part. Only in this way can the objectives of each form of test in a single examination be thoroughly planned in advance, and articulated so as to avoid unnecessary duplication and to assure proper adjustment of weights and consistent standards of rating.

This administrative plan does not mean that an examiner skilled in preparing and administering tests of knowledge in a particular occupational field is necessarily qualified to conduct a competitive oral test, even in his own field. Organizationally, the chief examiner can be held responsible for each part of the test and at the same time delegate to specialists the parts of the test they are qualified to handle. Thus, in an engineering examination, specialists in engineering may establish and apply standards for rating training and experience. But if the written test is divided into two parts, one dealing with knowledge of engineering techniques and one testing abstract intelligence, there is little in common between the two in so far as subject matter and approach are concerned. Whereas an engineering examiner would be quite competent to prepare the former, a specialist in psychological measurement would be needed for the latter.

Similarly, in developing the oral test program it is desirable, to the extent that the facilities of the personnel agency will permit, to use the knowledge of interview specialists—those who are fully conversant with the techniques, the dangers, and the potentialities of oral tests. In other words, the art of interviewing

is in itself a specialty. These considerations have led to the establishment of specialized interview sections within the examining divisions of some large agencies. In smaller agencies, it is desirable to have at least one qualified examiner who has general responsibility for the oral test program.

#### PLANNING FOR AN ORAL TEST

Among the fundamental problems involved in planning for oral tests are the selection of the factors to be rated by the examiners, and the determination of the standards to be employed. The task of acquainting the interviewers with the duties and requirements of the position, and the task of familiarizing them with the factors to be evidenced in the interview and the standards to be applied in rating are also a part of the preparation for the oral test and must be planned in advance.

The preliminary spadework consists in drawing up a detailed analysis of all possible duties of the class for which the examination is to be held, together with a description of all pertinent environmental conditions of employment and a statement of all the desirable qualifications. Such a job-description is a basic requirement for good planning. Although this analysis of duties and qualifications should be obtainable in most part from the class specifications, or perhaps in part from the announcement of the examination, it is unfortunately true that class specifications usually lack sufficient specific detail to provide examiners with the data needed in planning the examination. Such specifications often require considerable elaboration on the basis of conferences with appointing authorities and line supervisors.

When the necessary specifications are at hand, the several parts of the total examination, and the nature and scope of each, are determined: qualifying experience, written test, performance test, experience rating, character investigation, physical requirements, and oral test. When the position is one calling for an oral test, this test must be planned with due reference to its part in the total selective process. Wide differences of practice in the planning of oral tests are largely traceable to varying ideas as to the purpose and scope of these tests as related to the

other means used to compare the merits of applicants. To evaluate each candidate's education and experience, as well as his personal suitability for the position, is a duty assigned to oral examiners in some jurisdictions; in others the interviewers are instructed to avoid reference to the applicant's background and to concentrate on designated aspects of personality.

The relative merits of these different conceptions of the oral test have been discussed in Chapter IV. Here it need only be emphasized that clean-cut decisions regarding function and scope are the responsibility of the personnel administrator, not of the individual interviewers. Decisions regarding the ground to be covered, and the abilities, traits, capacities, dispositions, or other characteristics to be evaluated during the oral test must be precisely formulated before plans for the actual conduct of the test are made. In other words, determination of the qualifications to be evaluated should not be postponed until the oral test begins (as often occurs), or left to the sole judgment of the interviewers who happen to constitute the examining board. They should be predetermined and defined as part of a planned approach to the whole examination.

It is regrettable that in the rush of planning and carrying through a heavy schedule of examinations, there is a disposition on the part of many personnel agencies to defer so crucially important a task, and to leave it to the examiners as they confer shortly before the appearance of the first candidates. When the agency has an adequate, well-organized staff, any such side-stepping of responsibilities is inexcusable. Well in advance of the announcement of the examination its scope should be prescribed, the qualifications to be evaluated in each part should be concisely defined, and decisions should be reached as to the kinds of evidence to be sought and the standards of evaluation to be adopted. Preparation of the rating forms to be used in recording the examiners' evaluation is also a technical matter that calls for advance planning. Until these details have been completed it is impossible to proceed with the remaining steps in the planning of the oral test.

*Pre-Selection of Factors*

Much is to be said in favor of requiring that those who are to conduct the oral test confer at this initial stage with those responsible for selecting the factors to be rated. Qualified interviewers can help materially in clarifying and defining precisely the factors that can be tested best in the interview and in drawing attention to observable elements which should be considered in rating each factor. For example, if "alertness" is one such factor, what aspects of behavior do they regard as indicating that trait? Should the interviewer consider the candidate's rate of bodily movement, the tempo of his speech, the liveliness of his facial expression, or his mental readiness in repartee? Frequently it appears that failure to consider this or that observable component as relevant to a given factor is more responsible than anything else for disagreement among the raters.

In some jurisdictions the factors to be rated for all classes of positions have been standardized to a point where they are the same in all tests, and weights and rating standards are equally set by tradition. This situation strongly indicates that drastic overhauling is required in the examining division.

Again, in some places long-standing regulations require that factors be determined and standards be defined by the examiners alone. In others, the opinions of appointing officers as to the current requirements of the service are determinative. When the interviewers themselves cannot be consulted, there is no choice but to turn over the factors, as determined, to them, clarifying the reasons for the selection of the particular factors, and pointing out their relation to other parts of the test as fully as possible. But if the interviewers are available, and are as competent and well informed as they should be, it is to be expected that they will participate in the selection of the factors to be judged and in the formulation of standards.

When basic decisions underlying the interviewing procedure are arrived at in good time in the conference between the interviewers and other examiners and appointing officers, they will be far clearer in the minds of the interviewers, and there will be

greater assurance of wholehearted cooperation in carrying out the limited purpose of the oral test as planned than if such decisions were made at the last moment. Furthermore, the method here suggested makes full use of the intelligence and experience of the interviewers and lends itself to continuous improvement of procedure from examination to examination. Finally, this practice makes it possible to include in the examination announcement a statement describing the specific purposes and weights of each part of the oral test.

### *Planning the Rating Scheme*

Deciding upon the rating method and preparing the form on which the scale of values or the standards of fitness are defined, and on which the record of the interviewer's decisions may conveniently be made, is closely related to the decision concerning the factors to be considered. The form of the rating blank is, of course, largely determined by the factors and standards adopted. The factors or traits to be rated are usually listed and defined, but appropriate weights for each factor may or may not be assigned. If weights are assigned, a space for recording the examiner's numerical or letter rating is provided opposite each factor. A number of agencies take pains to prepare a useful graphic scale.<sup>1</sup>

Sometimes a traditional and static form is used. Preferably, the interviewers should be required to formulate a revised form for each new examination. In any event, it is essential that the form be so drawn up that the raters will have a clear idea of its meaning, and of the significance and the effect of the ratings they give. With reference to this latter point, it is necessary to inform the interviewers concerning two critical elements in the rating scheme: (1) That rating of a given factor, which, if the candidate were to stand or fall on that trait alone, would be regarded as barely passable. (2) That rating of the same factor which, since other compensating factors are to be rated, should result in the candidate's rejection on that ground, regardless of his ratings on all other traits.

<sup>1</sup> See Appendix B for specimens of oral test rating forms.



To illustrate, assume that an oral test is to be conducted for a position requiring constant contact with the public. Assume also that one of the defined factors to be rated is "courtesy." It is helpful to reach agreement first upon what constitutes the minimum evidence of courtesy sufficient to qualify if on other elements the candidate is also barely passing. Then, by contrast, there should be agreement on a somewhat lower minimum degree of courtesy which will qualify if there are compensatingly high ratings in other factors. The candidate may demonstrate a capacity to meet the public successfully in spite of a tendency toward brusqueness. On the other hand, evidences of arrogance may be such that for certain jobs he should not be passed in the interview at all, no matter how superior he may be in other particulars.

### *Preparation of Oral Test Material<sup>2</sup>*

The nature of the test material used, the topics to be discussed, and the questions to be asked are naturally derived from the factors to be evidenced and measured. The competitive principle demands that these be standardized as far as possible. If the number of candidates is small, uniformity in the topics and questions used can be achieved; but if the interviews extend over several days, then equivalent sets of topics and questions are desirable in order to prevent collusion among the candidates. In such a situation comparability as to scope and difficulty is the most that can be expected.

The foregoing discussion does not mean to imply that uniformity of questions and follow-up throughout the interview should always be sought. In fact, to arrive at a satisfactory basis for rating with equal thoroughness the same traits in all candidates, it is often necessary to vary the details of an interview to meet variations in the candidates' answers and reactions. But, for the sake of maintaining competitive conditions, basic topics and introductory questions should be the same as far as pos-

<sup>2</sup> Because planning inevitably involves a consideration of method and manner of administration, some discussion of these subjects will be included in this chapter. However, the limited references in this chapter should be used in relation to the fuller treatment of the same topics found in subsequent chapters.

sible, whatever minor variations circumstances may later impose or justify.

In testing for organizing and teaching positions, it may be desirable to test the candidate's ability to explain orally the underlying principles or philosophy of the job or of some of its special aspects. A useful form of test for this purpose is a selected paragraph containing some ideas on the subject, followed by a series of basic questions which serve to bring out the candidate's ability to explain, to illustrate, and to apply the substance of the excerpt, and to explain and defend his own opinion of the writer's views. Free-answer questions in the written test are often used to test comprehension of underlying principles, with about the same degree of rating difficulty. However, it is one thing to set forth ideas in writing and another to explain and maintain an opinion or a line of argument in a face-to-face discussion. The examiners must distinguish between the *content* of knowledge already tested, and the *presentation*, which is the ratable factor here.

Whatever the nature of the material used, it is common practice in many examining units to enlist the cooperation of the interviewers who are expert in the subject matter of the test in selecting the topics and the key questions. If more than one set of material is required, their knowledge as to the comparability of the content of the several sets is valuable. In any event, ample time should be allowed in advance of an oral test to permit the interviewers to become familiar with the requirements of the position, the factors to be rated, the material to be used, and the procedures to be followed. They must also be allowed the time necessary to reach a common basis of understanding concerning definitions and standards. Ways of accomplishing these ends are described in Chapter VI.

The preparation of the materials to be used in training interviewers is one of the time-consuming but essential steps in planning for an interview. The details of the oral test must be worked out early enough so that necessary material will be available to each examiner. The duplication and distribution of topics, questions, and scales should be surrounded with every precaution to prevent information leaking out.

*Scheduling the Tests*

The length of time necessary to conduct each interview depends upon the nature of the oral test and also upon the amount of time available. Arranging a time schedule for candidates to report for interviews is a simple matter. It is important that sufficient time be allowed between tests to permit the interviewers to discuss the evidence and to complete their records with care and deliberation. Ample time should also be allowed for each candidate, before the beginning of his test, to fill out whatever forms may be required, and to make any other preparation required by the test. In any event, ordinary courtesy requires that the candidate be allowed a few minutes of quiet after arriving to compose himself and to prepare for what he is sure to regard as an ordeal.

On the other hand, if candidates are required to wait unduly before being interviewed, they will suffer from prolonged anticipation. The practice of calling at one time all candidates to be examined during a forenoon or an afternoon is for this reason open to criticism. If the intention is to prevent collusion among candidates, this end is served far better by holding the first-called candidates after they have been interviewed until the last-called have arrived, and not permit them to confer with one another. If this cannot be done without causing dissatisfaction, a compromise solution may be achieved by ushering candidates out of the office by a separate exit after they have been interviewed.

No test schedule, however carefully devised, is of value unless it is adhered to, and it is often at this point that an oral test program breaks down. If the schedule is not observed strictly, the candidates called at the end of the list have reason to complain of nerve-trying delays. Interviewers often fail to keep interviews within the time limit agreed upon. Particularly is this the case in committee tests, for each member naturally wishes to participate in the questioning. To correct this tendency, it is advisable either to apportion the topics among the examiners in advance, or assign to one member of the board, or to each member in turn, the responsibility for conducting the

tests, with the understanding that he will allow time for supplementary questioning by the other members of the board. At the same time, the chairman should be vested with authority to terminate useless digressions. Needless to say, interruptions from outside are not to be tolerated. It may be helpful to place a clock, plainly visible to all, in the examination room. A member of the board or a proctor may then indicate by an unobtrusive signal that the test has only a few minutes more to go.

### *Plans for Proctoring*

It is customary to designate in advance of the test a staff of one or more proctors who are to receive the candidates, give them the necessary instructions, and keep the tests moving on schedule. If these proctors are not experienced in the work, it is desirable that they should receive their instructions in typewritten form and that these instructions be amplified in a brief conference. Moreover, it is a matter of some importance to select proctors who, by their cheerful disposition, their calm and courteous manner, and the sureness and general intelligence with which they perform their duties, will help to put the candidates at ease in advance of the tests.

### *Physical Preparations*

Fortunate is the personnel agency which has rooms well suited and equipped for conducting oral tests. Should ideal conditions not obtain, however, an effort should be made to achieve the best conditions possible. It is of primary importance that the examination room be free from distracting noises, either from the street or from adjacent rooms. Neither the candidates nor the examiners can be expected to do their best work in noisy surroundings. The wall should, if possible, be soundproof and acoustically treated. If this is the case they are also likely to be opaque, thus shutting out visual distractions as well as sounds. If artificial illumination is necessary, it should be suitable to the uses to which the room is to be put.

The furniture and its arrangement require forethought. To seat a candidate directly facing a row of examiners entrenched

behind a long table subjects him to unnecessary strain. A better arrangement is a circular table about which the interviewers and the candidate are to be seated, as for a conference; if a rectangular table is used, it should be small enough to permit the participants, examiners and candidate, to gather about it. The stenotypist or stenographer, if present, should be seated separately, near enough to hear the conversation, but not so near as to be a distraction.

Some personnel agencies have made effective use of sound recording devices to preserve a record of the interview. When this is done, the microphone should be made as inconspicuous as possible but candidates should be informed in advance that a record will be made. The recording apparatus and the technician are best located in another room. The controls for coordinating the interviewing and the recording should be operated by the interviewer rather than by the technician, that is, the former should indicate to the latter by signal the beginning and end of the interview. Finally, the advice of a sound-recording engineer should be followed with respect to sound-deadening treatment and other necessary arrangements in the examining room.

#### QUALIFICATIONS OF INTERVIEWERS

The perfect interviewer would be one of those rare beings who possess the best of human traits and qualities. In conferring with applicants for a variety of positions, an interviewer has to adjust himself to all kinds of personalities, under normal and unusual conditions. He must be ready to draw upon his own resources and energies to meet the unexpected and the unpredictable. He cannot be a slave to any system; he must adapt himself so completely to any situation that he may capture the full import of any remark, statement, expression, gesture, emotional reaction, or other clue to factual evidence desired. A good interviewer should be at least as intelligent as those he interviews—preferably more so. In a sense, the interview is a battle of wits in which the quicker wit should be on the side of the employer.

An interviewer must adjust his questioning to the mental capacities of the person interviewed. It is desirable that he possess a broad vocabulary which will enable him to express important shades of meaning and to understand replies made by candidates of wide and varied backgrounds.

### *Need for Job Knowledge*

A thorough knowledge of the duties and requirements of the position under discussion is imperative, but even though the interviewer may have a letter-perfect knowledge of the details of the job analysis and the requirements, this is not equivalent to actual work experience in the field. If he has never held such a position he is handicapped in interviewing an applicant who is thoroughly familiar with the work. But he can call in experts to advise him in preparing the interview. He can also add to his equipment by observing or working on the job. An interviewer who constantly explores the occupational world to increase his knowledge of a variety of jobs and of the qualifications of those who fill them satisfactorily will be much better qualified to do his own work.

A professionally competent interviewer has a working knowledge of the nature of evidence. He knows what kinds of information are relevant to his purpose, and what kinds are reliable and weighty. He has a logical mind and the ability to reason clearly.

### *Mental and Emotional Qualifications*

The mental and emotional equipment of the interviewer is of primary importance. He must possess a knowledge of the patterns of behavior and human motives which permits psychological understanding and sympathetic insight. An interview occasionally develops emotional tension. It is imperative that a good interviewer himself be emotionally stable. He should be neither stolid and phlegmatic nor flighty and neurotic. A well-poised examiner can best evaluate the flashes of emotional response which occur during interviews; he is not likely to be thrown off balance by temperamental displays. Compared with

these qualifications a fine appearance, dignified mien, and impressive stature, while advantageous, are relatively trivial.

Although it is obviously difficult to enumerate all of the qualifications of a good interviewer, a few of the more essential characteristics are included in the following list:<sup>3</sup>

1. Ability to inspire confidence in all kinds of people
2. Ability to be a good listener
3. Ability to put people at ease
4. Adaptability
5. Freedom from political, racial, and religious bias and personal prejudice
6. Familiarity with conditions of work in plants and offices, the surroundings, management, policies, and types of workers
7. Ability to keep the job specifications clearly in mind
8. Ability to talk the language of those interviewed
9. Analytical and critical ability
10. A kindly and sympathetic mien, combined with firmness and efficiency.
11. A sense of humor
12. An understanding of human motives
13. Ability to refrain from making impulsive judgments
14. A grasp of the nature of evidence and its probative value
15. Ability to keep adequate records
16. Readiness to criticize his own procedures after an interview

#### SELECTION OF INTERVIEWERS

A single individual, expert in interviewing, is sometimes assigned to conduct the oral test for reasons of economy. Some jurisdictions cannot afford to employ a number of interviewers and are unable to spare the time of busy examiners competent in a particular field to take part in the interviewing. For certain types of positions in which the factors to be tested are common, the procedure to be used is standardized, and a simple rating scale is established in advance, one expert may be com-

<sup>3</sup> Adapted from a section on interviewing in Jess T. Hopkins and Others, *The Emergence of a New Public Employment Service* (Public Employment Center of Rochester, 1935).

petent to conduct the tests. But allegations of unfairness, bias, or subjective rating are more likely to be leveled at a single examiner than at a board, and public confidence in the interviewing process may be shaken. Moreover, in examinations for specialized positions, when the factors to be tested include capacities related to the specialty, staff interviewers trained primarily in administering oral tests may not be competent to rate the value of the specialized evidence produced. In addition, an examining division will seldom have on its staff enough qualified interviewers who are also specialists in the many fields of activity for which it recruits.

For these reasons it is customary to set up interviewing boards with at least one trained interviewer and one or more additional examiners who are specialists in the occupational field involved. Sometimes, if outside "experts" are included, the boards are composed of as many as nine members. Such boards are able to combine the necessary varieties of specialized knowledges and the special techniques of oral testing.

### *Need for a Trained Chairman*

A trained and qualified interviewer should preside over the proceedings of the examining board. It is his duty to see that the discussion stays within the bounds of the allotted time and topic, and that relevant and reliable evidence is developed with reference to each factor to be rated. It is assumed, of course, that each member of the board will ask questions. Indeed, the particular function of the interviewer especially familiar with the occupational techniques in question is to ask whatever questions are necessary to bring out evidence of the quality of the candidates qualifications under each of the factors selected for rating. The chairman should see to it that the full value of all technical evidence is revealed. Continued questioning and polite but searching cross-examination, if necessary, is the duty of all members of the interviewing board. But the trained presiding interviewer, who leads the search for evidence and directs the deliberations, should guide the interview with a full knowledge of the dangers and limitations of the oral process.



*Departmental Representatives*

In some jurisdictions, departmental representatives are frequently called in to assist members of the oral board in conducting interviews for candidates for their particular departments. Other jurisdictions hold that impartiality may be lost when departmental representatives are permitted to share in the examining process. But, if the control of procedure remains properly in the hands of the representative of the central agency who presides over the board, there is much to be said in favor of using the services of departmental experts. They are likely to know best the conditions and requirements of the work for which the appointees are needed. They may be technically proficient in the type of work to be performed. Representation on the oral examining board tends to give the department greater confidence in the selective process, and appointing officers are more likely to be satisfied and to give their support to the merit system when they are participants.

To forestall any danger that departmental representatives may seek even unconsciously to aid favored applicants, the testing method must be carefully explained to them in advance, and the necessity for strict adherence to that method should be emphasized. It must be made clear what evidence will be considered relevant and reliable, as well as the procedure that may properly be used to adduce such evidence. The representative of the central personnel agency should direct the interview throughout and use the special knowledge and skill of the departmental representative to draw out and weigh evidence related to each factor. Finally, and most obviously, the standards of rating agreed to and approved by the personnel agency and the department in advance of the test must be strictly observed.

*Use of Outside Experts*

Outside experts are frequently employed to sit in on interviewing boards. Some advantages and disadvantages of this practice have been set forth in Chapter II. It should be emphasized that even though these examiners are eminent in their

own fields, they are not likely to know much about the methods and limitations of competitive civil service examinations, and even less about the fine points of oral testing. Time must be set aside for coaching the experts in advance of the test. They are of great assistance to the qualified interviewer in questioning the applicant to bring out the full quality and nature of his abilities. However, it is unwise to leave the conduct of the interview to them alone. Authorities differ on the question of whether such outside experts should be asked to rate the candidates.

### *Panels*

Panels of outside technicians have been selected in advance by some jurisdictions and given general training in the requirements of the interview process, so that they are prepared to serve from time to time as interviewers. These panels, from which special examiners are drawn as needed, may be established in widely diverse occupational fields so that the special knowledge needed to supplement the technical capacity of the personnel agency's own interviewers will always be available. Usually, however, the individuals selected from such panels to take part in a particular test require further detailed instruction about the duties and scope of the position, and about the definitions of factors, and the rating scale to be used.

### *Conclusions*

The number of interviewers on an oral testing board varies from one to nine. The larger the board, the more diverse may be the experience and skills of those who compose it, and the greater may be the variety of questions asked and the scope of evidence and "insight" obtained. The test may also expand far beyond its appointed and proper limits. When ratings have been based primarily on the judgment of the interviewers, it has been imagined that the greater the number of judges, and the longer the debates, the fairer the final approximation of fitness may be. But when only objective evidence is sought and a fixed

standard of rating that evidence is employed, there is little reason for the number of interviewers to be large.

It is the opinion of most administrators that the most serviceable board of interviewers includes the following members:<sup>4</sup>

1. A specialist in oral examining on the staff of the personnel agency who is trained and qualified to conduct the type of test ordered.

2. An expert in the field of occupation in which the test is being given.

3. A representative of the employing agency who is familiar with the conditions of work in his organization.

To train individual interviewers and to prepare the members of an oral board as a group to take part in a specific examination is an administrative responsibility for which planning is indispensable. In the following chapter attention will be focused on training methods and procedures that may be used to develop competent interviewers.

<sup>4</sup> When a larger board is constituted, it is not uncommon to select as additional examiners an employer or an experienced employment interviewer from industry, and an esteemed, public-spirited citizen who knows conditions in the community.

## *Chapter VI*

### **The Training of Interviewers**

**T**HE effectiveness of the oral test depends in great part on the skill of the examiners. Thus, the examiner should possess considerable aptitude for his job before he is selected to serve as a member of an examining board. Skill in the procedures of the interview, however, may be acquired and perfected through training. The successful personnel agency, therefore, steadfastly maintains high standards in its selection of interviewers and also systematically trains them after they have been selected. Some examiners are naturally more adroit than others in drawing out significant statements of fact and noting evidences of factors of fitness to be appraised, but the periods of training are valuable for all interviewers.

#### **NEED FOR PLANNED TRAINING**

For present purposes two classes of examiners may be distinguished: the professional and the novice. The professional interviewer is employed by a personnel agency to take charge of oral tests or to participate in them regularly as an official representative of the agency. His training must be more extensive and thorough than can be given to an outsider whose services as a special examiner are used only once or only occasionally. It must be recognized that a comprehensive program of training is possible only for interviewers who can devote the time required for continuously improving their methods.

The procuring of competent and objective interviewers who are willing to devote the time necessary for adequate preparation and performance of the task is of primary importance to the successful use of the oral test. But the substance and method of their training is of almost equal importance. Too often this preparation is done hastily during the moments immediately

preceding the appearance of the first candidate to be examined. Unhurried preparation has proved its worth in those agencies that give serious attention to the improvement of oral tests.<sup>1</sup>

### *Training the Novice*

A novice in oral testing is frequently chosen because he is an expert in a particular field. It is necessary, as a first step in his training, to make certain that he understands precisely the nature of the position for which the applicants are being interviewed. Sometimes it is sufficient that he review the published announcement of the examination. More frequently it is advisable to describe to him in greater detail the place of the position in the employing agency, the specific duties to be performed, and the working conditions. This description is commonly written and given to the interviewers before the test; but it should be reviewed and enlarged upon when the members of the board meet for final instructions. Only after all the members have grasped thoroughly the duties and responsibilities of the position should their attention be turned to the next point: the qualifications deemed necessary, however measurable, in order to perform these duties satisfactorily. The examining agency will already have defined these qualifications and described them succinctly in its published announcement of the examination. However, it is essential that they subsequently be discussed in detail with the less experienced interviewers to make certain that there is no misunderstanding.

The interviewers need to be reminded also that not all these qualifications, however desirable or necessary they may be, can be ascertained satisfactorily during the course of an oral test. Some of them are brought to light and measured with greater precision by means of the written test, by independent appraisal

<sup>1</sup> Louis J. Kroeger, former Executive Officer of the California State Personnel Board, says: "We have attempted to standardize our practices and standards of judgment in each examination, first, by having a general discussion of standards to be observed in each test conducted before any of the candidates are interviewed, and secondly, by having the member of the qualifications appraisal board who is a member of our staff (and who serves as chairman of the board) travel from one part of the state to another, even though the other members of the qualifications appraisal board in each locality are generally drawn from the community in an effort to reduce expense."

of the applicant's experience record, or by a subsequent character investigation. It is therefore customary to describe to the examiner the successive steps in the total process of recruitment and selection, so that he will appreciate the unique function of the oral test as only one step in the process and be prepared to concentrate attention on the evidences of fitness that can best be secured through that medium.

### INSTRUCTION MANUALS

A carefully prepared manual or written statement covering the aims, policies, and procedures of the examining agency is a most useful device in instructing novices. A description of the position for which the examination is to be held and statements regarding the qualifications sought and the factors to be assessed should also be available for study.

For these purposes the Tennessee Valley Authority uses a 10-page "Interviewer's Manual," supplemented by more extensive and detailed manuals of instruction covering specific topics. For example, the "Interviewer's Rating Scale for Trades and Labor Classification" distills for the inexperienced interviewer the results of extensive experience in examining for these positions and in training interviewers for this work.

A "Manual of Suggestions for Oral Examining Boards," issued by the Employment Board of the Department of Public Assistance in Pennsylvania, in March, 1938, contains within its 60 pages the following documents and descriptions:

Letters to Members of Oral Examining Boards

Function of Oral Examinations (with description of the process of recruitment, examination, certification, and appointment)

Composition of Oral Boards

Oral Examinations in General

Aims of Examinations

Organization of Oral Boards

Suggestions for General Conduct of Interview

The Questioning Process (with examples of bad procedure as well as good)

The Rating Process

In an appendix are found sample materials: salary scales for the positions to be filled; job descriptions, including scope of job, nature of work and personal qualifications necessary for its performance; questions suggested for use in the test; and a sample rating form, together with instructions for its use.

In a section of the manual devoted to oral examinations in general, the policy that has been adopted by this agency is made explicit:<sup>2</sup>

This particular set of examinations is planned to grade certain personal characteristics vitally important in the work of the Department which can be brought to light in an interview but which cannot be measured in a written examination nor in a grading of past experience. The following facts about these oral interviews should be borne in mind:

1. The purpose of the oral examination should not be to put candidates "on the spot," but rather to give them an opportunity to put themselves across in the most favorable light.

2. The purpose of the interviews is not to rate training and experience. That is being done in another part of the examination.

3. The purpose of oral examinations is not to secure information on the applicant's familiarity with work processes and the subject matter of the job. This information has been tested by the written examination.

4. The oral examinations should not be used to attempt to measure such personality factors as honesty, courage, and unselfishness on which evidence cannot be produced and accurately judged in an oral interview.

5. No attempt should be made during the oral examination to discover political, religious, labor, or fraternal affiliations or sympathies of the applicant.

6. The oral examinations should not be used in lieu of a character investigation or to test the accuracy of previously made statements, or to get opinions of former employers.

7. The oral examinations are not to take the place of a physical examination which may be arranged for later.

8. The oral examination does not take the place of the employment interview which is conducted by the prospective employer in order to decide who will be selected from the certified list.

<sup>2</sup> Other agencies follow an essentially different policy regarding the scope of the oral test, as the reader of the following pages will recognize.

The "Aim" is stated in part in these words:

To make the whole enterprise completely honest, impartial, and fair to all concerned . . . so that each applicant on leaving the oral interview shall be confident that he has been given every reasonable opportunity.

Under "Organization of Oral Boards," it is stated that:

Each oral board . . . should have a secretary and a chairman. . . . The function of the secretary is to be responsible for the custody of the records and the completeness of the rating sheets. The function of the chairman is to see that the interview is properly conducted.

The specific duties of the secretary and of the chairman are then outlined in detail.

The Civil Service Commission of San Francisco uses a concise two-page instruction sheet called "Suggestions and General Information to Special Examiners in Rating Candidates in Oral Tests." The place of the oral test in the total process of examining is briefly indicated. The items to be graded are described as follows:

1. Personal Characteristics
  - a. Appearance, bearing, poise, and manner
  - b. Penetration and powers of comprehension
  - c. Address, thought, and speech
2. Personal History
  - a. Education
  - b. Training and experience
3. General Fitness to Hold Position

From this instruction sheet a few injunctions and sample questions may be quoted:

Do not rate the candidate on his technical knowledge or lack of knowledge on the job. This is measured by a written examination. Base your estimates of the applicant's characteristics solely on evidences observed during the interview.

After you have interviewed several candidates, you are at liberty to revise your ratings. You may discuss the personal suitability of the candidates with the other examiners, but this is not advised. It is recommended that each examiner record his ratings on the basis



of his independent judgment. The ratings of the three examiners will be averaged by the staff of the Civil Service Commission and these general averages only will be shown to the candidate. Your rating sheet and the markings thereon are not subject to review by the candidate.

The questions to be asked by examiners will relate mostly to the candidate's personal history, i.e., education, and training and experience.

The examiners are not expected to follow a prescribed line of questioning, but are requested to make the interview informal and to follow up any significant leads which the candidate's remarks suggest. Each candidate must, however, be asked each of the following basic questions:

1. What, in your conception, are the more important features of this position?
2. What training and experience have you had that you believe qualifies you to hold this position?
3. What have you learned from your past experience that would be helpful to you in carrying out the responsibilities of this position?
4. What education have you had which in your judgment fits you for this position?

The examiners will have available the applications of the candidates and from the experience statement shown on the application, questions may be asked. Before the examination, a short meeting of the special examiners will be held to explain any questions in the minds of the special examiners regarding the conduct of the oral examination.

The items under Personal Characteristics and General Fitness can be noted by the candidate's responses to these questions. It is suggested that a predetermined leadership be arranged for each interview. Other members of the board should pick up the conversation before it lags. See that most of the talking is done by the applicant. Make the range of questions wide enough to bring out behavior indicative of each of the characteristics to be appraised.

In connection with a nation-wide unassembled examination for the position of Child-Welfare Consultant, the Arkansas State Personnel Division developed a special manual, "Oral Interview Procedure and Instructions," for study by members of the several boards which conducted interviews in various cities. The following topics are discussed: Purpose of Interview, Duties and Qualifications, Plan of Interview, Conduct of

Interview, Length of Interview, Professional and Technical Phase, Questions and Responses, Recording Responses, Reactions of Applicants, Closing of Interview, Personality Rating Summary Report. The paragraph under the heading "Professional and Technical Phase" is indicative at once of the scope of this oral test and of the kind of instructions given to the examiners:

The first part of the interview is intended to act as a guide in determining the scope of the applicant's experience and professional judgment. A series of general discussion questions has been prepared, each followed by a check-list of salient points, all or part of which should be covered in the response of a person who is thoroughly familiar with the subject. The interviewer is to cover all questions with each applicant, indicating on the check-list those points covered in the applicant's response. Any other responses, not covered in any way in the check-list, which appear to the interviewer to be pertinent, should be indicated in essential detail in the space below the check-list. The interviewer must use his judgment here, and include only such additional points as appear to represent an original, pertinent thought.

The Los Angeles County Commission's "Manual of Civil Service Procedure" deals with the oral test under the following headings:<sup>3</sup>

1. Arranging
  - a. In setting date of oral interview, consider examination schedule, total number of applicants, number of applicants likely to be eliminated in written examinations, etc.
  - b. If interview is to be held before written papers are marked, decide date to permit notices of oral interview to be distributed to applicants on day of written examination.
  - c. Confer with Secretary of Examining Division to arrange schedule for interviews. Determine time to be allotted for each interview, total number of applicants to be interviewed each day, etc.

<sup>3</sup> Clifford N. Amsden, Chief Examiner of the Los Angeles County Civil Service Commission, makes this statement: "Outside persons serving on qualifications appraisal boards are instructed orally by the personnel examiner in charge of the examination. The position is discussed at length in terms of the requirements and the factors to be rated in the oral interview. The procedure to be followed in conducting the interview and any special interview report forms or rating schedules are considered and explained."

2. Planning
  - a. Outline general procedure to be followed in each interview, points to be covered, etc. If possible, devise rating schedule to be used in evaluating experience, education, training, and general qualifications of applicants.
  - b. In planning interviews for an examination, consider material obtained in preliminary work on examination, relating to duties and requirements of the position and general qualifications required of applicants.
3. Conducting
  - a. Observe following general rules in conducting oral interviews:
    - (1) Allow adequate time for each interview.
    - (2) Maintain uniform conditions, in so far as possible, for all applicants.
    - (3) Use any procedures and devices which will allow maximum objectivity in rating applications.
    - (4) Establish rapport with each applicant and try to put him at ease. Give applicant every advantage in order to obtain complete information relating to experience, education, training, and general qualifications, upon which to base evaluation.
  - b. On oral sheets or special forms provided for this purpose, make extensive notations during each interview concerning applicant's education, experience, training, skills, and other qualifications for the position involved.
  - c. Record rating given to each applicant on oral sheets or special forms.
  - d. Indicate by check marks on applications the names of former employers to whom Reference Forms should be sent by Record Section.
  - e. Send all applications, oral sheets, and any other forms used to Record Section for computation of final averages.

Other agencies which have successfully approached the task of training interviewers have embodied their statements of policy and practice in manuals for study, or in detailed letters containing information which the examiners can assimilate before coming together to receive additional instruction.

While there is substantial agreement among most of these examining bodies as to their aims and procedures in oral tests, striking differences are occasionally observed. Some civil serv-

ice commissions advocate separating sharply the grading of the candidate's technical knowledge, ascertained in the written examination, from the appraisal of his personal suitability, ascertained in the interview.<sup>4</sup>

Whatever the local practice may be, the examiner should be left in no uncertainty regarding the underlying philosophy and the current procedure of the examining agency he represents.

A training manual for oral examiners should also define the standards and specifications to be observed, should describe the personal characteristics, traits, abilities, or other factors to be appraised, and should explain the kinds of evidence which indicate the degree to which the candidates possess these characteristics. Sample questions and topics of discussion that have proved their suitability as means of drawing out such evidences should be listed, together with suggestions as to how to escape the pitfalls of improper questioning and irrelevant discussion. The study of such manuals or letters should be supplemented by instruction given in conferences and discussion meetings.

### *Instruction in the Total Selection Process*

It is evident that the subject matter to be covered during the training of interviewers should begin with a description of the entire process of merit system examination and selection. Interviewers perform their part of the total task with greater effectiveness when they understand the purposes of the selective

<sup>4</sup>Dr. Bingham submits the following discussion of practice: The theory seems to be that a candidate's knowledge is measured by means of written examinations more adequately than by oral tests, and that this factor should not be rated by two or three different instruments of measurement. In other jurisdictions the opposite view is held, namely, that additional information regarding the candidate's technical knowledge, obtainable during the oral test, is a valuable supplement to the information gained in the written examination, and that consequently efforts of interviewers to sample the candidate's technical knowledge and proficiency are not inappropriate in the oral examination. These agencies maintain that every valid means of ascertaining technical competence should be employed. They also hold that evidence of ability to deal crisply and judiciously with technical questions in the interview, as well as evidence of ability to write about such topics, is important. It has even been alleged that it is impossible adequately to sample the evidences of a candidate's personal suitability for the position without observing the competence with which he deals with technical matters during the interview.

procedure and the part played in it by the written test, the rating of training and experience as recorded on the application form, the character investigation, the medical and physical tests, and the oral test.

The process of combining the data from these several sources to establish the rank order of the candidates on an eligible list need not be elaborated in all its details, but the instructor should be prepared to answer any questions about these procedures. The process of certification to be followed by the personnel agency should also be described. If the employing agency is to have the privilege of conducting its own employment interviews and selecting employees from among the eligibles certified, this point should be made clear. At some point in this orientation process it should also be stressed that the purpose of the selective procedure is not only to make certain that the candidates are correctly and fairly ranked in the order of their relative fitness. It is also necessary to eliminate from the list all who are clearly unsuited, since it may happen that every one who receives a passing mark will have to be employed before the list expires.

Interviewers need to be reminded that, in the interest of fairness to all competitors, it is essential to adhere strictly to the rules and regulations established by the examining agency. Every step in the process of selection is subject to review; hence the necessity for adequate records, not only of the ratings given by the examiners but also of the evidence on which these appraisals are made.

A description of the entire process followed by the particular agency, and of the merit system philosophy that underlies it, should be available for study. When the whole process of recruitment, selection, certification, and employment is comprehended, it becomes easier to grasp the proper function of the oral test within this total process. It becomes obvious that an oral test serves chiefly to appraise those factors of relative fitness that cannot be measured more accurately, if at all, by means of written tests and experience ratings, namely, the candidate's qualifications as evidenced by his past behavior and ac-

accomplishments under relevant circumstances and his behavior during the interview.

It should be made clear that only in this test can full evidence of desirable and undesirable personal factors be developed, and that in the oral process there may be observed the bases for the personal impression each candidate makes on others, his ability to grasp directly the point of a question and to discuss it intelligently, his command of language, the clarity and correctness of his speech, his ability to talk to the point instead of wandering, and his ability to discuss with competence and discernment matters of importance within his special field. His emotional stability under circumstances of the examination may be observed, together with his physical bearing. If the duties which the candidate may be called upon to undertake involve a good deal of personal contact with subordinates, associates, or the public, his behavior during the interview will be especially significant.

The interviewers should be given a thorough understanding of the specific factors, traits, or characteristics to be observed and rated in the interview. It is not enough to name general factors such as presence, tact, initiative, ability to cooperate, convincingness, personal suitability. Each factor should be broken down into components clearly and succinctly defined in advance, reduced to writing, and studied and discussed with the interviewers, so that there may be agreement among them as to what is to be rated and what evidence is to be sought.

It is necessary to emphasize that these characteristics are not to be appraised in the abstract but as factors in the total fitness of the person to do the work of the particular position. Thus "ingenuity" among competitors for a position as a detective or a field investigator may differ in quality and definition from the "ingenuity" desired in an office supervisor, a construction engineer, or a hospital superintendent. At no time during the training of the interviewer or during his service as an examiner should he be permitted to forget that he is assisting to ascertain the relative abilities of competitors for appointment to a certain class of position.

We have seen that not all examining agencies have the same concept of the function and scope of the oral test. Some agencies limit it strictly to consideration of factors not ascertainable by written examination or experience ratings; others place a major part of the burden of selection on the oral test. Thus, the interviewers may be instructed to explore every lead for evidences of technical competence, quality of experience, and educational background as well as personal suitability.

For all these reasons, therefore, the interviewers should be informed as to the philosophy and policies of the agency, and the scope of the oral test in the particular examination in which they are to participate.

#### TRAINING IN THE TECHNIQUES OF ORAL TESTING

Having established a clear understanding regarding the policies of the examining agency, having defined the duties of the position and the qualifications deemed essential, and having agreed on the nature and meaning of the factors to be rated, it is next in order for the examiner to discuss the methods by which the desired evidence is elicited during the interview. Sample questions and topics for discussion with the applicant may be reviewed. Questions that may bring forth information about political affiliations, race, or religious preferences should be avoided. The desirability of keeping the interview informal and spontaneous should be emphasized; nevertheless, before it has ended, sufficient evidence should have been elicited to provide a basis for rating each trait or factor to be appraised.

#### *The Need for Objectivity*

The interviewer should again be reminded that the test will be subject to review if any competitor chooses to appeal. Each rating, therefore, must not be based upon subjective impression but upon reviewable evidence—objective evidence obtained during the course of the examination, evidence clearly relevant to the competitor's suitability for appointment, and sufficient to furnish a reliable basis of appraisal.

The distinction between evidence and inference cannot be

too strongly emphasized.<sup>5</sup> Indeed, one of the most difficult points to make clear to interviewers seems to be this difference between the securing of evidence and the evaluation or appraisal of this evidence. There is a strong tendency to telescope these processes. Analysis of evidence has reached its highest development in the trial procedures of law courts which have long been concerned with fair methods of proof. While oral examiners are not confined by legalistic rules of proof, their task is quite analogous to that of a court. Both seek to arrive at a just conclusion from a set of facts. An oral examining board is at once questioner, jury, and judge. The correctness of its decisions is a result of the skill with which pertinent significant evidence is elicited and the objectivity with which this evidence is analyzed and evaluated as an indication of the relative fitness for the work to be done by the appointee.<sup>6</sup> Practice in recognizing evidence is obviously an essential part of the interviewer's training. No less essential is practice in evaluating its significance.

Interviewers are helped to achieve greater reliability in their judgments by drill in distinguishing between objective evidence and subjective impression. Also, when evaluating objective evidence they should be reminded that all humans are subject to personal bias of one kind or another. They need to recognize and discount any personal prejudices or preferences which they

<sup>5</sup> See Samuel H. Ordway, Jr., and James C. O'Brien, *An Approach to More Objective Oral Tests* (Washington, D. C.: Society for Personnel Administration, Pamphlet No. 2, June 1939). This analysis of the problem, with a description of a tried procedure for securing in the oral test objective evidence which may be independently evaluated, has an immediate use in the training of interviewers to appreciate the intricacies of their task and to follow sound procedures in performing it. Those who are responsible for planning oral tests and for training interviewers will find here a method of questioning definitely designed to draw out pertinent, objective, reliable evidence regarding factors of performance (what the candidate has actually done, and what he actually does during the interview) which can then be evaluated as objective evidence of the kind and quality of performance to be expected on the job. A revised description of this approach is found in Chapter IV. The pamphlet describes a single approach, however, and should not be offered as the only useful method of securing evidence.

<sup>6</sup> Separation of the functions of gathering evidence and evaluating evidence is illustrated in the employee rating procedures followed by the Civil Service Commissions of the City of New York, Dallas, Los Angeles, and a number of other jurisdictions. See Samuel H. Ordway, Jr., and John C. Laffan, *Approaches to the Measurement and Reward of Effective Work of Individual Government Employees* (New York: National Municipal League, published as a supplement to the *National Municipal Review*, October 1935).



may be harboring—for or against fat men, red-headed women, or economic royalists—which otherwise tend to vitiate our appraisals of the evidence they present.

### *Instruction in Rating Procedure*

Next it is appropriate for the instructor to describe the ways prescribed for recording examiners' ratings and to answer questions about the graphic rating form or the numerical scale to be used in expressing the judgments of the examiners. A little practice in using any such device is certain to provoke several questions. Must the examiners reach agreement as to their ratings? If they are to record their ratings independently, are they permitted to compare notes and to discuss with each other the evidence, like a jury? Is it permissible for an examiner to revise his ratings after several competitors have been interviewed? What is the examiner to do if he finds that the evidence regarding some one factor seems to him wholly insufficient as a basis for appraisal? Can the candidate then be recalled for further questioning? Are examiners expected to note or consider any evidences they may observe of personal characteristics not specifically mentioned on the rating form but which nevertheless might be a distinct handicap or an asset to the candidate in his relations with subordinates, associates, and public? To these and similar inquiries the instructor or the chairman of the board of examiners must have authoritative replies.

In reaching an agreement regarding standards of comparison, the interviewers are certain to raise questions as to the passing mark. They want to know what minimum evidences of each factor may reasonably be expected of a candidate who is qualified for the position in question. Few instructors are able to isolate this critical point on the scale of evaluation to their complete satisfaction.

Different jurisdictions meet the challenge in various ways. In some the examiners are informed that a rating of 60 is passing (or 70, or 50, or "D" on a scale of letter grades), and that the examining agency defers to the interviewers' judgment as to the evidence which will warrant giving a "D" or "passing mark"

or a rating of "satisfactory" on each of the several traits to be rated. They are told, for instance, that the degree worthy of a rating of 60 on "appearance" is a matter of individual judgment with each examiner. In other jurisdictions "60" is defined as "60 per cent." "Per cent of what?" the interviewer asks. "Do you mean an amount of this factor which is lacking in the least suitable 60 per cent of the general population? Or the least suitable 60 per cent of the applicants? Or of employees? Or do you mean 60 per cent of the amount of this factor found in the perfect 100 per cent person? Or is '100' a rating applicable, not to an ideal person, but to a candidate who resembles the most satisfactory employee actually filling this kind of position?"

The instructor is prone to accept this last interpretation as a way out of the dilemma, and to hasten on to a discussion of less precarious topics. The important thing to make clear, however, is that the interviewers must always affirmatively ascertain the standard which the examining agency wishes them to apply and then determine the kinds and amounts of evidence they will accept to meet that standard for each step on the scale they are asked to use.

There should always be discussion and ultimate agreement among the examiners as to the kinds of evidence they will recognize as indicating (1) barely acceptable qualifications, and (2) superb qualifications for a certain position. In this way the low and high points on the scale may be fixed. Disagreement on these points among examiners should not be reconciled by argument. Instead, two or three practice interviews should be held, and the prescribed rating technique should be put to use. The attention of the interviewers is thus focused on the actual behavior of persons considered as candidates competing for a definite job.

#### *Agreement on General Policy and Procedure*

Details of procedure during the oral test must also be agreed upon. Who will assume leadership in the interview? Which of the examiners will open the discussion of the specific points to be covered? How can the atmosphere of the interview be kept

informal and friendly? How much time will be made available? Which of the examiners will be responsible for terminating the interview, and how should this be done?

The steps following the interview should also be made clear to answer such questions as these: Is discussion of the evidence encouraged, or are the examiners required to evaluate their findings independently without consultation? In the event a candidate is disqualified, is it required that a written record of this verdict be stated in detail? May an unsupported general conclusion be submitted? If not, what kinds of reasons are deemed valid? What kinds of evidence of fitness may be expected to stand the scrutiny of subsequent review in the event the candidate appeals?

### *Value of Practice Interviews*

Practice in interviewing pseudo candidates is one of the most effective and rewarding devices for familiarizing examiners with standard procedures. This training device brings out and puts the proper emphasis on all points of procedure and technique. The practice interview affords the examiners an opportunity to gather skill in questioning and in noting significant behavior. Forms of questioning that are inappropriate soon come to be recognized; the more important observations to be made are impressed upon them. They gain confidence in making judgments and in using the prescribed method of rating.

Subsequent to each practice interview there is an opportunity to review and discuss the candidate's qualifications in the light of the facts disclosed. Superfluous questions are noted; the need for further questioning in certain directions becomes obvious; familiarity with the use of the rating form is required. Most important of all, the practice interview permits members of the board to establish standards of comparison. In this connection, they will be materially aided in clarifying their scale of values if they are given opportunity to interview at least three candidates, one of whom would be barely acceptable as an employee, one of whom would be quite acceptable, and one of whom would be a keen competitor for top rank on the eligible list.

*Mechanical Training Aids*

Mechanical aids in the training of interviewers have included the use of motion pictures and phonographic recordings. Recordings of the examiner's voice and emphasis when asking questions, and of his gestures, appearance, and manner while interviewing, have been tried as aids in drawing attention to details needing correction. The same mechanical devices have also been employed in presenting to interviewers the records of what candidates have said and how they have behaved during an oral test, thus serving the same training function as the practice interview.

*Some Common Mistakes*

During the discussions which follow these practice interviews, the examiners may be forewarned of the most common pitfalls.<sup>7</sup> Attention should be drawn, for example, to the tendency of some interviewers to let their general impression of the candidate warp their judgment of specific traits. This weakness, which may exhibit itself in various ways, such as a tendency to rate an obviously unsuitable candidate equally low on all traits, can be largely eradicated with practice. A similar tendency is to make use of only a small part of the available scale of values—to bunch the ratings of most of the candidates about the average, or close to the passing mark, or near the top of the scale. It can be corrected by reminding the novice that sharp discrimination is desirable, though forced distribution must be avoided. It is also desirable to stress that, unless ratings are fairly well distributed along the scale of values, they will necessarily have little weight in determining the competitors' relative positions when averaged with the marks given by other examiners and with standings in the written examination.<sup>8</sup>

This particular pitfall, curiously enough, is encountered more frequently in rating easily observable traits like "speaking

<sup>7</sup> See Chapter VII for a more detailed discussion of this subject.

<sup>8</sup> An illustration of the effect of range of ratings on rank order of candidates is given in W. V. Bingham, *Oral Examinations in Civil Service Recruitment* (Chicago: Civil Service Assembly, Pamphlet No. 13, February 1939).

voice," "appearance," and "command of language" than in making judgments about more complex and obscure traits such as "ability to plan and organize." There is generally closer agreement among examiners regarding degrees of variation in a candidate's "ability to plan and organize" and similar relatively complex traits than there is with reference to such apparently objective and readily observed traits as "appearance" and "command of language." This finding may be due in part to the fact that evidences of abilities are more factual and less impressionistic than evidences of traits, and also to the fact that examiners give less careful attention to the rating of apparently obvious traits. It may also be due in part to a tendency on the part of raters, in considering these obvious traits, to reflect their own likes and dislikes, and to leave out of account the bearing of these traits on the suitability of the candidate for the particular position specified. Most frequently, however, it is due to lack of definition of the component parts of these obvious traits.

Interviewers cannot be reminded too emphatically to keep constantly before them the work an appointee will be called upon to perform. It is not a candidate's "appearance in general" or his "command of language in all situations" which is being evaluated, but rather his "appearance in so far as it will be an asset or a liability to him in the work of *policeman*," or his "command of language needed in the work of a *mediator of labor disputes*." The tendency to rate factors in the abstract, neglecting their relevance to the particular job to be undertaken, is a common fault that training and experience will correct.

In summary, the need for training of interviewers arises partly from the necessity of establishing and maintaining uniform standards of procedure and of evaluation. The content of such training has here been briefly outlined. Other sections of this survey deal with these topics in greater detail. When these points have been formulated in a manual of suggestions for interviewers, a necessary first step has been taken toward their training.

We have seen that it is not sufficient to hand such a manual

to an examiner for individual study. Personal conferences which emphasize and elaborate the more essential topics are also necessary. Practice in interviewing is then suggested, with subsequent opportunity for further conference regarding ideal procedures and pitfalls to be avoided.

### TRAINING THE PROFESSIONAL INTERVIEWER

For the professional interviewer, similar review and discussion of various types of interview, and of the process and results of the interviews he has completed, with comparison of the methods which he and others have used, contribute to his further training and increase his ability to bring out necessary evidence of relative fitness and to place a just evaluation on this evidence.

The development of competence in a professional interviewer is a task that requires more varied, elaborate, and thorough training than that of a novice. The professional interviewer must be prepared to train less experienced examiners and supervise their work, as well as to take part himself in the planning and conduct of oral tests. This training should be a never-ending process. It should include understanding of the duties and requirements of the widest possible range of occupations.

#### *Scope of Professional Training*

The broad objective of training is to develop wisdom and skill in oral testing by equipping the professional interviewer with the requisite knowledge of the positions for which applicants are to compete, the qualifications sought, the standards established, and the procedures that are used to measure the fitness of the competitors. The professional interviewer must be fully acquainted with all steps in the selective process, from the writing of class specifications and preparation of defined qualifications standards to the planning and conduct of total examinations, the certification of eligibles, the hearing of appeals, and the follow-up on appointees during and after the probationary period. Provision should be made for him to observe the actual work of each of these steps and to take part in most of them.

### *Assignments for Experience*

In some of the more forward-looking jurisdictions, the neophyte oral examiner receives a planned series of assignments, each of which contributes to his knowledge and experience. He takes part in staff conferences in which the basic requirements of the various positions for which examinations are scheduled are debated and defined. He is assigned as an observer to the rating of training and experience. In this way he absorbs an appreciation of the types of eligibles sought, the nature of the openings to be filled, and the kinds of performance deemed to be indicative of the qualifications desired.

The apprentice interviewer also helps to define and to reduce to writing the standards of rating to be applied in forthcoming examinations and in the adjudication of any appeals that may be taken from the ratings given in those examinations. He serves in the certification division and learns the demands and objections of appointing officers. He sits in at conferences in which the head of the oral testing unit and a representative of the board of review deal with appeals from oral ratings. The broader his experience, the clearer becomes his perspective and the greater his appreciation of the consequences of unskillful interviewing, hasty rating, and inadequate recording. He is supplied with available studies in the field of oral testing and with copies of court opinions concerning oral tests. He is not left to absorb these materials unaided, but is given opportunity to confer with his superiors and associates, and to consider and weigh the material in terms of the usefulness to the work of the agency.

This system of training makes for a thorough understanding of the aims and objectives of the oral testing process and tends to create an *esprit de corps* indispensable to administration of any program so subject to change and continuous improvement.

### ILLUSTRATIONS OF TRAINING PROGRAMS

#### *Field Investigators*

In many respects the training of oral examiners resembles the training of field investigators whose duty it is to conduct char-

acter investigations or to secure information about candidates by conferring with people in a position to know about them. The procedures developed by the United States Civil Service Commission for training their field examiners (investigators) are thus worthy of consideration as a pattern by personnel administrators responsible for training interviewers, no matter what their specific assignments are to be.

The duties of these field examiners are not limited to interviewing of candidates to verify experience and to report obvious evidences of personal unfitness for the position. Former employers and associates must also be interviewed in order to check the accuracy of data supplied by the applicant and to gather additional facts bearing upon his character and suitability for government service. The training has the twofold objective of increasing the interviewer's proficiency in securing the necessary factual evidence, and teaching him to write reports that will stand the scrutiny of critical review and furnish a substantial basis upon which others may determine the candidate's suitability. While these investigators do not rate qualifications, they are subject to many of the limitations and restrictions of rating examiners.

*Initial Assignments.* The Commission provides a *Field Examiners Manual of Procedure* of 121 pages to give the investigators explicit instructions and materials for study. Training begins with observing the work of a seasoned investigator who is his teacher for the time being. Then he is permitted to take the lead in interviews, his instructor accompanying him and, when necessary, supplementing his inquiries until the trainee is judged to be competent to carry on alone. Preliminary to any specific assignment the interviewer is given a clear picture of the duties, responsibilities, and requirements of the position for which the candidate is being tested, for example, a superintendent of a veteran's hospital, a border-patrol inspector, or a postmaster. The necessity of verifying with special care the evidence relative to work record, character, and personal history facts is stressed.

*Training Conferences.* Problems that have arisen in the course of the observed tests are subsequently discussed in daily



meetings of trainees and experienced investigators. Questions are raised; constructive suggestions are offered. Reports of interviews are read and the evidence is weighed in the light of all the information they contain. Sample reports of personal interviews in the Appendix of the *Field Examiner's Manual* are discussed and compared with reports currently received, so that the trainees may acquire skill in the preparation of succinct and illuminating reports as well as in securing significant data. Practice in reviewing the evidence contained in reports prepared by other interviewers impresses the trainee with the necessity for full and accurate recording of essential data, and helps him to improve his own reports. Special attention is given to training him in the technique of dictating a report.

After some months of experience in the field, the interviewer is transferred for a few weeks to headquarters, where he gains further experience in reviewing the evidence found in the reports of other investigators. This experience tends to improve his own reports when he subsequently returns to the field.

During the entire period of training every encouragement is given the trainee. His good work is commended—his ingenuity in searching out and securing evidence, his adroitness and tact in questioning, his progress in preparing full and accurate reports and in estimating the worth of the data they embody. He learns to avoid mistakes and to surmount difficulties by following and observing the work of seasoned interviewers.

### *Other Programs*

The training programs for professional interviewers in other agencies include experience in securing from representatives of the employing bureaus the data with which to prepare job descriptions and specifications, in drafting announcements of examinations, in scheduling and arranging for these examinations, and in presiding over oral tests in which they act as examiners. The interviewers acquire familiarity with the policies and regulations of the agency and with the nature of the positions for which oral tests are conducted by helping other members of the staff with their duties—participating, for example, in the

work of following up the efficiency ratings of employees, or investigating complaints. The more systematically this variety of experience is scheduled, the more promptly the trainee gains the desired background.

Some interviewers then specialize in planning and conducting examinations for a certain field—for skilled trades, for engineering occupations, for professional fields, for clerical positions, or for supervisory and administrative posts. They learn all they can about the duties and requirements of positions in this one field, and the reasons for the relative success or failure of employees in these posts. Courses of reading or of advanced academic study in government administration, employment management, vocational psychology, and related subjects are sometimes prescribed; familiarity with contributions appearing in current books and journals is recognized as desirable.

Of undoubted worth in the training of interviewers is participation as opportunity offers in meetings of professional associations like the Civil Service Assembly, the Society for Personnel Administration, and the Personnel Division of the American Management Association.

### *Self-Training and Research*

An interviewer can learn much in reviewing his own successes and mistakes. After an oral test or a personal investigation has been completed, a short period of reflection brings into perspective those steps which have brought the most satisfactory results and the questions which drew out the most significant replies. Any errors or omissions that become apparent may suggest that a different technique be tried in the next interview. Did the interviewer waste precious time in a roundabout attempt to establish desired rapport, or did he get off to a prompt start? Did he let the interview get out of hand or keep it to the point? Did he see that all essential questions were covered? Were the facts brought to light sufficient on which to base the ratings? What additional questions might well have been asked? These and similar questions indicate the desirable line which self-analysis should take.

Such an appraisal of his procedure while the interview is still fresh in his mind is certain to suggest to the interviewer ways in which errors may be avoided and more adequate data secured. He may discover the need for more thorough preparation, for a clearer understanding of the precise duties of the positions, or for advance formulation of the wording of key questions. The interviewer may similarly review his ratings and his report in an effort to improve his work. This kind of scrutiny will help to develop proficiency and confidence in his procedures.

### *Research as a Training Aid*

Another avenue of self-development is opened whenever the interviewer can take part in research and experiment. Fruitful topics of inquiry are legion. There is scarcely a section of this report that does not suggest to the thoughtful reader a question that calls for further investigation.<sup>9</sup> To cite a single example, an interviewer who can participate in a study designed to check upon the subsequent success or failure of appointees, and to establish definite relationships between an objective measure of success on the job and the measures of probable success which were recorded at the time of the oral test, cannot fail to improve thereby his own grasp of the problems confronting every interviewer.

<sup>9</sup> For additional suggestions concerning desirable lines of research in the field of oral testing, see Leonard D. White, *Research in Public Personnel Administration* (Washington, D.C.: Social Science Research Council, Committee on Public Administration, 1942), p. 13.

## *Chapter VII*

### **Oral Test Procedures**

THE seasoned interviewer's attitude toward the candidate during an interview is kindly and encouraging, and marked by a certain firmness which should command (but not dominate) the first impression. Friendliness facilitates the proceedings and firmness prevents waste of time. Some of the more valuable observations to be obtained during an interview depend on the spontaneous and uninhibited reaction of the candidate. Data developed in free discussion are often more objective than considered answers to leading questions, but free response cannot be obtained if the interviewer assumes a totally impersonal and frozen manner. An understanding smile, a nod of appreciation encourages the applicant. If he feels that in the interviewer he has a friend at court, he is apt to give information more freely. The interviewer's goal should be to draw forth the whole evidence in a businesslike and friendly way.

#### **THE EXAMINER'S MANNER AND BEARING**

The posture of the interviewer has been the subject of several enlightening experiments. Their results indicate that if the interviewer assumes an erect position in his chair, thus indicating that the occasion is businesslike and slightly formal, the person interviewed is apt to be less discursive and more direct and terse in his replies. A slouchy, carefree, or indolent position has the opposite effect. The upright, square-shouldered, straight-back position gives the impression that time is valuable. Savings of 10 to 15 per cent in the time involved in interviews have been demonstrated when the businesslike posture was used, in contrast to the slouchy or carefree, "feet-on-the-desk" position. When the correct position is coupled with good diction and a well-modulated voice, time is usually saved and a good impression is made on the candidate.

### *Identification of Candidates*

In some interviews it is first necessary to determine whether the person who appears is the candidate whose name has been presented. A photograph or fingerprint may serve the purpose of identification; identity may be verified by competent procuring before the interview begins; or a signature may be requested. If there is a suspicion of fraud, it may be necessary to call upon the investigative staff. But the interview should not be begun in an atmosphere of obvious suspicion.

### CONDUCTING THE INTERVIEW

An unaffected greeting and handshake go far to establish rapport and to put the candidate at ease, even though he is prone to be nervous and under a decided strain. He should always be told the length and purpose of the interview. The interview should open with the discussion of a topic of mutual interest. The best and most obvious one is the nature of the job for which the candidate is applying, and his special qualifications for filling it. Too often interviewers waste precious moments beating about the bush, commenting on the unusual weather or some other irrelevancy in order to establish rapport; a more productive way to insure rapport is to proceed at once to the topic of primary interest.

Some short explanation may be needed as to the purpose of the conference or of the method to be followed. If the plan does not require regimented questioning of the candidates, each may be asked to tell freely about some phase of his past experience which illustrates his ability to fill the position.

### *Questions and Discourse*

The questions asked should be phrased clearly and in the simplest possible terms. Except as they have been planned in advance and are to be directed to all candidates, they should be injected only when it is necessary to direct the course of the interview to an issue, to bring out information which is being overlooked, or to check for additional data. Even when specific

questions have been formulated before the interview, their exact wording is sometimes varied to suit the personality and background of each candidate. It is often helpful to rephrase in a different manner a question once asked to bring out additional information or to check on the reliability of a former statement. Heckling should never be permitted, but if the candidate is suspected of falsifying, he should be cross-examined vigorously. If he is too discursive in his replies, direct questions, not heckling, are likely to elicit the necessary information. Questions that suggest that the interviewer expects any particular answer should be avoided at all cost.

The interviewer is interested in developing evidence about required qualifications; he therefore should have a definite object in each inquiry he makes during the interview. But, at least in the early stages of the test, it is worth while to let the applicant express himself freely in his own way, provided he is not too verbose or discursive. In so doing he reveals a good deal of information concerning his attitudes and personality, which may be observed and noted by the interviewer. Habits of thinking and expression are vividly illustrated when the candidate is permitted to talk without interruption. The easy flow of free discussion must not be permitted, however, to turn into a rambling discussion. When a candidate strays entirely off the subject, the interviewer can bring him back with a pertinent question. If he continues to digress, the interviewer may resort to the direct question-and-answer technique.

### Notes

Good notes are indispensable for the record. Even when the entire interview is recorded by a stenographer the verbatim report will not include the facts regarding the applicant's gestures and mannerisms, which are often significant. If the interview is not recorded verbatim, *constant* notetaking by the examiners is essential for an adequate record. Taking notes during an interview may sometimes be distracting to the applicant, but ordinarily he is not bothered if the notetaking is unobtrusive. Continual notetaking bothers the candidate far less than occa-

sional jotting, which gives him the impression that he has made some misstatement or has been guilty of some *faux pas*. No pains should be spared to get as complete a record as possible of all significant information brought out during the interview. Even material which seems insignificant at the time may prove important later.

### *Timing the Interview*

When many candidates are to be interviewed, reasonable adherence to the time schedule is imperative. The necessity for obtaining all essential, pertinent information and at the same time adhering to a time schedule scrupulously points directly to the value of skillful planning and management of the interview. The reputation of an organization suffers from defamatory reports by impatient candidates who must wait long after the hour at which their interviews have been scheduled. When the interview discloses lack of requisite qualifying experience or other evidence which definitely eliminates the candidate, there is no use in prolonging the conference; it should be courteously terminated. The candidate who cannot possibly qualify will then not entertain false hopes.

### *Termination*

In terminating the interview it should be remembered that one of the purposes of the organization is to build up good will. A moment of relaxation, humor, or at least friendly conversation at the close of the ordeal is appreciated by the candidate regardless of the success with which he has met the oral test.

Immediately after the interview, the pertinent evidence should be reviewed and summarized. If this process is postponed until after subsequent interviews, confusion is inevitable. Observations and ratings should be carefully recorded while the candidate's performance is fresh in mind. This does not preclude standardized revision of the ratings later, after several candidates have been interviewed and after a final standard has been established for applying the rating scale to the evidence obtained in the interview.

## HAZARDS AND PITFALLS TO BE AVOIDED

Throughout the oral test and at its close, there are a number of common hazards that must be foreseen and avoided, lest the outcome of the entire interview be impaired. Some of these concern the attitudes of the interviewer; others involve his reasoning processes in reviewing and appraising the evidence; still others have to do with the artifices, conscious or otherwise, whereby the candidate endeavors to sway the interviewer's judgment in his favor. Awareness of the existence of these pitfalls is essential to sound oral testing procedure, for only by identifying them can the interviewer take steps to circumvent them.

*Pseudo-Scientific Methods*

Nearly everyone thinks he is a good judge of human nature. Indeed, the belief in one's inherent capacity to size up the other fellow is almost universal. This is one of the reasons why the interviewing process is so popular in employment procedure.

From time immemorial man has attempted to foretell the future and to devise short-cut methods for judging his fellow-men. Systems for reading character have been prevalent in all ages. It is estimated that a large part of the population still believes in the pseudo sciences of phrenology and physiognomy. Astrology, palmistry, phrenology, physiognomy, numerology, graphology, and character analysis (a modern name for the systems of physiognomy and phrenology which first flourished a century ago) all have some adherents in business and even in government agencies. Many scientifically controlled experiments have demonstrated that all of these pseudo-scientific attempts to read character are worthless, and none of them has any standing in scientific circles.

It is not appropriate here to attempt a full exposition of the claims of those who promote such systems. All who are charged with the responsibility for interviewing candidates for positions should be familiar with these claims and should be in possession of the facts that refute them. So widespread are these "gold-brick" methods that their terminology (blond traits, the convex



type, etc.) has crept into our language. The fact that they sometimes offer easy explanations for the possession or lack of a character trait wins support for them. Ordinarily these systems are built on the observation of a few striking cases, and rarely if ever have their proponents attempted to offer statistical proof of the validity of their claims.

So much have the terms and beliefs of the pseudo sciences become a part of daily life that the interviewer must be on constant guard lest he be unconsciously affected by them. It is so easy to think that a person with a weak handshake may have a weak character, or that one with eyes set close together will be narrow-minded, that a square chin indicates dominance, a long sharp nose inquisitiveness, a high forehead a brilliant intellect, piercing eyes shrewdness, and so on. All of these and similar claims have been thoroughly disproven. Suffice it to say that there is no positive correlation between physical features and mental and emotional capacities. The interviewer who allows his judgment to be colored by such folklore and by the abracadabra of occult professors thereby confesses his reluctance or inability to approach a set of facts rationally and objectively. The way to find out how a man's central nervous system functions is by talking with him, giving him tests, and observing his behavior—not by noting the cut of his jaw or the shape of his cranium.

This does not imply that appearance has no relationship to success on certain jobs. Of course a neat appearance, a pleasant smile, good looks, and similar characteristics are sometimes assets in kinds of work where the individual is brought in close contact with others. But these qualifications will be rarely conclusive.

### *Conditioned Reactions and "Halo"*

Much that we know about others and the effect they have upon us is based on our past experience. Thus, in meeting a person for the first time, we are immediately impressed, consciously or unconsciously, by something he does or says which reminds us of someone else, or some past train of experience.

Our reactions toward new situations and individuals are based on our own past. If a rabbit is placed with a small child, the child will ordinarily pet it. But John Watson long ago showed experimentally that if a loud noise is made behind the child when the rabbit is present, the frightened child associates the loud noise with the rabbit and henceforth is afraid of rabbits. This fear is a conditioned reaction. Much of our learning and most of our fears and emotions have their origin in this conditioning process.

Conditioned reactions thus account for the first impressions an interviewer gets of individuals as he interviews them. Something about their manner or approach, their stride, gestures, voice, or appearance resembles someone else he has known. He tends to feel toward this new individual the same emotions and attitudes he has toward the person or situation they bring to mind. If the stranger who is to be interviewed reminds the interviewer of someone he dislikes, his attitude toward the candidate is biased against him; but if the candidate reminds him of a friend he is apt to be prejudiced in his favor. It is difficult to maintain a thoroughly objective mental attitude, but this state must be achieved by the interviewer. Obviously, it is unfair to attribute traits and characteristics to an individual simply because he reminds the interviewer of someone else. Since there is no absolutely established relationship between physical and mental or emotional traits, there is no basis for assuming that, because a person's external appearance reminds us of another, he therefore possesses similar characteristics.

Frequently the interviewer's general reaction to a candidate is very favorable and leads him to give the candidate an equally high rating on all factors. Something in his behavior has made an excellent impression and it is very human to tend to give him credit for possessing all or many of the better personality traits. Or he may have impressed the interviewer with his mediocrity so that the tendency is to rate him "average" in all traits, without proper discrimination. The tendency to let a general impression color judgment of specific traits is the "halo" weakness, familiar to most interviewers. To safeguard against halo it is

necessary to focus attention separately on each characteristic or sub-trait in relation to the requirements of the job itself. An interviewer who has been drilled in using a well-prepared rating scale is not likely to be trapped by this pitfall.<sup>1</sup>

### *The Habit of Generalizing*

Almost everyone is influenced at times by the phenomenon known in psychology as the "generalizing habit." We tend to think that if a person shows a trait in one situation it will carry over into other situations. Thus it is assumed that a person who is neat in his dress will do neat paper work, or that an individual who responds readily in speech will be equally quick in his work motions. But who has not observed the workman who turns out a neat piece of handicraft; yet who works in the midst of apparent confusion? Or the man who makes a fetish of neatness at the expense of output?

This tendency toward generalization can often be very misleading. Thus, if an interviewer is faced with two applicants for a clerical position, one of whom is dressed in rough work clothes and the other in a conventional business suit, there is a natural tendency to assume that the latter would make the better clerk. Yet questioning may reveal that it was necessary for the first candidate to come to the interview directly from his job on a production line. He may in fact have the better qualifications of the two, and it is a mistake to allow superficial appearance to detract from his rating. It may even be that the circumstances indicate an asset of character in that he was willing to accept another kind of work while awaiting an opportunity to get back into the clerical work in which he was most proficient.

Most habits are specific; that is, they are pertinent only to particular circumstances. It is not safe to assume that because an individual has been scrupulously honest in money matters he will also be honest on the golf course. True, good traits are ordinarily found together; but it is not safe to assume that indi-

<sup>1</sup> There are some interviewing situations in which halo is believed to be inevitable and to be valid. See W. V. Bingham, "Halo, Invalid and Valid," *Journal of Applied Psychology*, April 1939.

cation of the trait brought out in an interview will surely be demonstrated in work situations. Furthermore, many persons are under a strain during an interview and their behavior is not typical. This tension may show itself in unaccustomed alertness or timidity, in anger or haughtiness which would not be assumed by them under ordinary circumstances. It is well for the interviewer to keep watch on his own mental processes during and after the interview to make sure that he does not indulge in the so-called "generalizing habit."

Since the object of any interview is to obtain objective information concerning another person, anything done by the interviewer that may color this information should be avoided. Emotions and attitudes are contagious. If pictures of a sour-visaged individual and a smiling individual are successively held before a group of people there would be no difficulty at all in knowing which picture the group was being shown. The persons in the group would reflect the picture in their facial expressions. This is known as instinctive imitation or the imitative tendency. The happy and cheerful individual generates that feeling in those he interviews; the grumpy and pessimistic individual affects others in the same way. The careful interviewer guards his own feelings. He realizes that the impressions he receives depend in part on how he himself feels, and that the attitudes he assumes are picked up unconsciously by those who come before him. Many a competent applicant for a job in industry has failed to be hired for no better reason than that the employment interviewer was suffering from insomnia or had an argument prior to the interview. Anger and fear reactions are particularly contagious. This can be seen in the way a crowd or mob becomes panicky when one or two individuals lose their heads or show terror.

### *The Use of Suggestion*

Suggestion, either in the manner in which the question is asked, or in citing analogies or examples, must be used with utmost caution. Probably it should never be used in the form of a question by means of which the interviewer seeks to get direct information. Occasionally it may be legitimate to suggest

a line of thought to the candidate which serves as a "primer," provided that no specific answer to the question is suggested. The interviewer's duty is to direct and systemize the interview without injecting himself into it by suggesting any desired replies. If, however, the candidate finds difficulty in understanding the question, it is permissible to give examples not too closely resembling the anticipated answer. This is only a form of priming to expedite the interview.

Of course, suggestibility is a trait which may deserve evaluation for its own sake. The person who doesn't easily rise to a lead is at a disadvantage in some forms of employment, such as teaching. If the candidate's suggestibility is to be rated, the interviewer must offer him some stimulus. The point is that suggestion on the part of the interviewer should not be used unless he has a reason for it and is aware of what he is doing.

### *Wishful Thinking*

Almost all applicants for a position are eager to please the interviewer and to present themselves in the best possible light. With this cast of mind, some exaggeration is to be expected in the effort to make their experience appear to the best advantage; for example, they commonly overstress their skills. Even though the exaggeration may be slight in the answer to any single question, the cumulative effect of many such answers may be considerable. The industrial specialists who interviewed drafted men during the first World War reported that less than 6 per cent actually possessed the kind and degree of skill they claimed. The tendency to overrate capacity through imagined qualifications must be watched for in all employment interviews.

Occasionally an individual underrates himself. When properly used, the art of understatement produces a favorable effect on an interviewer. If he discovers that the candidate has more experience or ability than he claims, the favorable effect is apt to be strong. To evaluate experience which seems to be overstated or understated the interviewer must seek detailed data. Specific questions should be asked until he is satisfied that he has ample accurate information.

*The Struggle to Dominate*

Some interviewers are under the false impression that they should completely dominate the situation; the result is antagonism. Antagonism in turn, arouses various emotional responses in both parties to the interview. Fear, defensiveness, apprehension, hatred—all by-products of antagonism—have the effect of destroying the objectivity of an interview. The word “interview” itself implies an exchange of views and does not in any way connote dominance.

*Fatigue*

Excessive fatigue on the part of either the interviewer or the candidate affects the situation adversely. Responses are not so quickly or easily made and the mind is not so alert under the condition of fatigue. Experiments show that there is a loss of attention and an increased tendency to distraction when either party to the interview is tired. Interviews should be so scheduled and managed that this condition will not arise. If they do, it is the responsibility of the interviewer to relieve the situation if possible.

A competitive interview reaches its climax when the predetermined standards of fitness are applied to the evidence, and ratings are assigned. Chapter VIII deals with the problems faced by oral examiners when rating candidates, and the importance of preserving an accurate record of the evidences brought out and observed during the interview.

## Chapter VIII

### Rating and Recording

THE act of rating involves at least three steps: analyzing or assorting the evidence to be rated; evaluating the evidence by applying to it defined standards of value; and computing and recording the result. For each factor and for many combinations of factors, where evidences are judged according to standards of excellence, the rating is commonly expressed on a numerical or alphabetical scale, a graphic linear scale, or some other prescribed measuring rod. The rating forms in current use are of various kinds. They are alike, however, in that they are all designed to aid interviewers in making evaluations and translating them into quantitative terms.<sup>1</sup>

#### DEVELOPING THE RATING PROCEDURE

The form in which the examiner's valuations are recorded—graphic ratings, numerical ratings, increments or decrements, percentages, or letter grades—should be determined in advance of the test. If graphic scales are used, the steps on each scale should be defined in well-chosen, descriptive phrases to identify the extremes of the scale, the midpoint, and at least two intermediate steps. In recent years the tendency in civil service agencies has been to avoid defining more than five—or at most seven—steps on such a graphic scale. If a rater wants to make finer discriminations than this, as he sometimes does, there is no reason why he should not indicate them by checking at intermediate points. Indeed, it is desirable to encourage examiners to make finer distinctions by asking them to review their ratings after they have examined several candidates. An interviewer may then find that he has checked three candidates at the same

<sup>1</sup> Examples of rating scales in current use appear in Appendix B of this report. The construction of rating scales and forms for use in various kinds of oral tests has been well set forth in *Manual of Merit System Administration*, prepared by the Social Security Board.

point on the scale, and yet he may be certain that the evidence shows that Candidate B is, after all, a little better than Candidate A and not quite as good as Candidate C, on the particular factor under consideration, although he would not wish to change any rating as much as would be implied by moving it a whole step of the scale.<sup>2</sup>

### *Factors and Over-All Rating*

Characteristics readily observed, such as speaking voice, neatness of dress, command of language, and poise, are ordinarily segregated, either first or last on the rating form, if they are to be rated at all. Factors which are evidenced in different ways, behavior patterns on the one hand, or demonstrated capacities, such as initiative or ability to gain cooperation, on the other, are likewise segregated. An over-all rating of the candidate's fitness for the position, if such a rating is required in spite of all the objections to that practice discussed in this report, is usually placed at the bottom of the sheet by way of summary.

Although evidence applicable to any one or more factors may be observed or noted at any time during the interview, rating is facilitated by considering separately all the evidence related to one factor at a time: all the evidence (usually observations) of neatness, or all the examples of demonstration of initiative, or of ability to supervise, or all the evidences which together may disclose a particular behavior pattern, or whatever other ratable component of fitness is in order. Indeed, in conducting the interview, notation should be made on the rating scale under the appropriate factor of each related item of evidence as it is disclosed; but the interviewer should seek so far as possible to develop at one time evidence related to each factor. Concentration upon developing the fullest possible evidence about each component in schematic order is a valuable aid to thoroughness. As soon as the examiner has secured, by observation and dis-

<sup>2</sup> It has occasionally been said that raters should not be permitted to record their judgments in units smaller than a known discriminable difference. Although the procedure here discussed may tend to convey a spurious impression of accuracy, no statistical harm results.—*W. F. B.*



cussion, sufficient evidence to rate the applicant fairly on one factor, he concentrates attention on the next factor on the list.

The ability to give his attention to the several components in turn is of great advantage to the interviewer. Only in this way is he likely to avoid having his general impression influence his final rating on any specific factor. He can thus isolate and weigh all the evidence relevant to each separate factor and can marshal that combination of facts which is known to establish certain patterns of conduct. In this way application of standards becomes more precise.

### *Internal Consistency of Ratings*

It has been repeatedly shown that over-all estimates of personal fitness are likely to be more consistent and valid if the interviewer has first carefully evaluated all the evidence pertaining to each specific factor. This is only natural if the factors have probative value of fitness, for a composite rating, based on specific weighted ratings on all factors, should be more accurate than uncontrolled over-all judgment.

Sometimes an inexperienced examiner is disconcerted to find that his own over-all estimate of the applicant's qualifications does not exactly correspond to the final rating awarded the applicant on the basis of the specific ratings of the factors of personal fitness. He may even be tempted to readjust his ratings of the factors to make them conform to his judgment in this respect. It may be that the evidence needs to be reconsidered, but the rater should first remember that over-all fitness for a specified job is something other than the correct evaluation of the few qualifications tested in the interview. Experience and knowledge of duties, not properly ratable factors in the interview, may have entered into his own over-all judgment. It is the applicant's experience and knowledge, plus his abilities and aptitude, plus his emotional adjustment, plus his energy and health, plus his mental adjustment, plus every other one of the factors of fitness tested at some part of the over-all examinations which make up the ultimate estimate of fitness.

A person is more than the sum of his separate organs, abilities, and qualities. A candidate who has been given at least a passing mark on each separate trait rated in the interview may nevertheless fail to convince the rater that he is fit to fill the position. Or, on the other hand, the total impression of the person's fitness may be so well balanced, may seem so well adapted to the requirements of the position to be filled, that the over-all impression is genuinely more favorable than the average of ratings on the several specific interview factors.

Lack of understanding as to the precise and limited part of the interview in the total process may explain the erroneous reasoning of some jurisdictions which require the members of an oral board to add their own over-all evaluations of each candidate, and give to that over-all rating a weight which goes far to offset the more objective, weighted ratings on each separate interview factor. It may also explain why other jurisdictions, knowing that various other elements of fitness have been rated in other tests, prohibit over-all ratings in the interview, lest such addenda unbalance a carefully scheduled array of weights for different parts of the examination. It should be noted that there is little uniformity of procedure in this matter of including a final over-all "factor."

### *Independent Rating vs. Consultations*

In some agencies, interviewers are required to arrive at their ratings independently of one another, without consultation. In others, they may be required to compare and reach approximate agreement in their separate ratings, or to agree on them within specified limits of variation. Much is to be said for the practice of calling first for independent ratings. This practice avoids the danger that one member of the board may dominate it. Thus, it is well to require each interviewer to weigh for himself the evidence on each ratable factor. But after this has been done, something of value is gained by subsequent comparison of ratings and joint discussion of the evidence. If the raters find that they are in wide disagreement, opportunity may then be

given for all to review the evidence and for any interviewer to revise his ratings in the event he finds he has overlooked any significant evidence.

### *Time of Rating*

The suggestion that assignment of ratings be deferred until after all candidates have been interviewed has merit only if relatively few applicants are to be examined and if they are not to be ranked on the basis of any extrinsic standard. When the volume of competition exceeds a bare minimum, preliminary ratings should be assigned at the conclusion of each interview. Even when a full and reviewable record of the interview is preserved, it is a time-consuming process for any examiner or board of examiners at some later time to go over the whole record again with sufficient care to evaluate the evidence developed concerning each factor for each candidate in the interview. Deferred rating, in the absence of such a careful review of the record, tends to a rationalization of worth, by the assignment of values to some particular evidence that is recalled, rather than to a critical analysis of all the evidence received.

### *Rating by Amateurs*

The common practice of inviting local citizens or civic leaders to sit on oral boards has already been described. When examining boards are so constituted, it is highly undesirable that the untrained and inexperienced interviewers be given the authority to assign final oral ratings. The public interest and the rights, welfare, hopes, and ambitions of individuals are as much involved in oral tests as in other competitive procedures. Standardization, difficult enough for professionals to maintain, is essential. For this reason, if for no other, the scoring of oral tests by novices should be avoided. When arrangements have been made for nontechnical or lay representation on oral examining boards, it should be understood at the outset that these representative citizens are present in the capacity of advisors and observers. They may be asked to take part in the questioning, but

when the final rating is assigned, their function should be advisory. If outside specialists are members of the board they may be called upon to rate factors within their special cognizance.<sup>3</sup>

### *Predetermination of Weighting Standards*

It should not be within the province of the interviewers to decide the weight to be given each factor except within the limits set by the examining division. Generally, the weight of each part of an examination, and of each factor in a part, should be predetermined by those in charge of the whole examining process. The relative importance of interdependent elements of fitness is one of the most difficult determinations to make.<sup>4</sup> Research is sorely needed to develop and establish more flexible yet standardized methods of weighting that will yield under various conditions the closest approximation to the ultimate criterion: success on the job. But weights should not be subject to adjusting at the pleasure of interviewers in order to bring ratings closer to their subjective estimates of the applicants.

### *Machine Averaging*

When large numbers of candidates compete, some use has been made of machine averaging of the ratings given by the oral examiners. If several factors having different weights are independently rated, the speed and accuracy of the machines in computing final weighted averages cuts down clerical costs and shortens the time required to make available to the employing agencies the lists of certified applicants.<sup>5</sup>

## THE ORAL TEST RECORD

Whether procedures for the review of the oral process are formalized or not, it is essential for internal discipline as well as

<sup>3</sup> Louis J. Kroeger observes that if the procedures recommended in this report are followed for (1) developing carefully the factors to be rated, (2) determining the relative weights to be given each factor, (3) selecting representatives of outside interests for the value of their contribution and for good public relations, and (4) training these representatives in method and standards, then it is a mistake not to let them participate equally in the rating.

<sup>4</sup> This subject has been covered in some detail in Chapter IV.

<sup>5</sup> See Reuben Horchow, *Machines in Civil Service Recruitment* (Chicago: Civil Service Assembly, Pamphlet No. 14, October 1939).

for fairness to the competitors that there be made at each interview a record stating in full the procedure followed, the factors used, the evidence adduced, the basis of rating, and the ratings resulting from the interview. Practice and method of making a record of interviews varies greatly in different jurisdictions. The principal possibilities will be listed and discussed briefly.

### *Kinds of Records*

*Longhand Notes.* The simplest form of record and that most frequently used is longhand notation of significant statements by applicants, together with observations made by the examiners. This record is usually followed by an opinion setting forth the reasons why a certain rating has been given. The use of the written opinion, however, has decreased in recent years. Such opinions tend to degenerate into subjective statements of the examiners' unsupported conclusions which are properly open to attack by disgruntled applicants.

*Stenographic or Stenotype Record.* It is generally agreed that a complete stenographic or stenotype record of the questions and answers in the course of the interview, plus a statement of factual observations, is a better form of record for review than longhand notes. At the close of each interview the examiners should promptly supplement the verbatim record with an objective description of what they have observed including all pertinent facts that would not appear as part of the spoken record but that may have some bearing upon the rating awarded. The stenotypist's record need not be transcribed unless there is need for review.

*Sound Recordings.* A number of jurisdictions today have improved upon stenographic records by installing one of a number of mechanical systems for recording questions and answers on phonographic records. The microphone used for such recording is inconspicuously placed so as not to distract the attention of examiners or applicants. The applicant should be told in advance that a recording of the interview will be made. This system has the advantage not only of recording the voices and emphases, but also of indicating pauses and hesitations in the

answers. A complete record of the spoken evidence is thus preserved and the examiners, of course, are also free at the end to record whatever factual observations may be significant in the rating.

*Sound Motion Picture.* Probably the ideal way to record an interview is by making a sound motion picture. Advantages of this method are that it would provide a permanent record not only of each question asked, each answer given, and the emphasis placed on questions and answers, but also of the facial expressions, diction, tone, gestures, and mannerisms of the interviewers and the applicants. But the unnatural conditions under which such a record would have to be made and the expense involved would rule out this method in most jurisdictions.

### *Recording Observations*

While dress, mannerisms, and gestures are not automatically recordable by any mechanical means short of photography, they often constitute an important part of the evidence which should be preserved for review. The examiners' statements of what they have observed become therefore a significant part of the record and examiners should be trained to make their statements as objective, as accurate, and as factual as possible. "This candidate has two gold front teeth and a scar which twists the left side of his mouth out of line" would be such a statement. A statement such as "The applicant was a disagreeable individual, with unattractive appearance and poor personality" is unsupported by recorded evidence, and does not constitute a statement of factual observation.<sup>6</sup> On the basis of facts observed, an

<sup>6</sup> *Comments from committee correspondence:* "However, a statement that 'this applicant impressed the interviewer as being a pleasant individual . . . ' is a statement of fact which would be relevant if the impression the candidates make is to be rated. A board of review would then properly take it into account. The fact that the applicant made this impression on an experienced interviewer is some evidence that he would impress other people similarly."—*W. V. B.*

As any lawyer will confirm, the statement of impression quoted is not a fact but a conclusion. The quotation, even if a statement of correct conclusion, is unreviewable because unsupported by recorded evidence of the observations on which the impression was founded. I agree that the fact the expert examiner was of the opinion he was impressed will impress some other impressionable people, and that it is *some* evidence; better evidence would be the facts and observations which were the grounds for the examiner's impression, if, indeed, it was grounded in fact.

examiner might be justified in concluding that there should be a low rating of this candidate on appearance and personality, but it is the facts, not his conclusions, that should appear on the record.

Thus a general guide to the recording of factual data lies in the answers to several questions: What facts impressed the examiner with reference to one candidate, which he did not observe or which differed from the facts he observed in another? What specific observation enabled him to conclude that Candidate A was superior to Candidate B in manner, personality, or appearance—if such general factors have a weightable value in the test? In this connection a factual statement might be: "This candidate sat erect and, without fidgeting, looked his questioner straight in the eye when answering; an occasional smile, particularly before answering questions 2 and 4 indicated his appreciation of their nicety."

Such observations may not be, and usually are not by themselves, sufficient evidence on which to base a final rating on any particular component of fitness. But many similar observations, taken together with other data, may disclose a pattern of conduct or traits of ratable significance. It is the composite of many relevant items that becomes significant. But it is the items that may become significant which should be recorded.

### *Use of the Record for Review*

A primary use of the record is to clear up, at the time the examiners are arriving at their rating, possible differences in recollections as to what the candidate answered or how he behaved. When no record has been preserved, differences in recollections or opinions among examiners are usually resolved by compromise. With a record immediately available for review, the examiners can resolve the difference on a factual basis. Moreover, the record is frequently useful when a candidate who feels that he has not received a sufficiently high rating comes in to discuss his rating with the examiners. An adequate record of the interview should explain the basis for the conclusions reached by the examiners in the final rating.

In jurisdictions where a formal procedure has been provided whereby appeals can be made either to a board of appeals, to a commission itself, or to a single administrator, or where review in the courts is authorized by law, a full record of the oral test is indispensable. To review a case fairly, an appellate body must have before it the evidence upon which the rating was given and the evidence on which ratings were given to other candidates in the same test. Although appellate bodies will not ordinarily substitute their judgment for that of the examiners in any matter involving the weight of the evidence, they are likely to order a new examination in the absence of any evidence relevant to a rated factor of fitness, or in the event a rating has been based on evidence wholly irrelevant to the factor in question.

With an adequate record of the interview given to each applicant, it should not be necessary for any examining agency to have to resort to a re-test of an aggrieved individual unless there occurred demonstrable errors. Of course, a record is not a guarantee against error. Obviously, if the test is shown to be unfair in some other particular a re-test may be necessary despite a complete record. For example, it may be that a candidate who is exhausted and excited by excessive waiting at the time of his test should in fairness be entitled to re-examination. But this fact is extrinsic to the record of the interview itself.

The practice in some jurisdictions of arranging for a second interview or a re-test, in the case of an aggrieved individual whose protest is sufficiently vigorous, is unsound if the original test was valid and an adequate record of the basis on which the rating was awarded has been preserved. The oral process, if properly recorded, should be just as reviewable as any other part of the examining technique.

### *The Danger of Excessive Formalism*

In developing a good record system for the purposes of review it is necessary to avoid one common danger: the making of a legalistic record which actually does not preserve the basis of rating. This danger is well expressed in the following statement by Dr. Arthur W. Kornhauser:



Isolated bits of evidence (gestures, expressions, reported incidents) which can be exactly specified in the record are likely *not* to be significant bases of our estimates of a candidate's qualifications. It is the composite picture which is revealing. Each fragment is interpreted as it fits into the gradually evolving size-up of the person. Its meaning depends upon its consistency with other data about the individual. To set down the separate observations and then mechanically to rate them gives the appearance of objectivity and may satisfy a court, but has little promise, in my opinion, of achieving genuinely valuable appraisals of candidates. It usually results, I believe, in spurious and self-deceptive accuracy and objectivity as far as the real purpose of the rating is concerned.

Simple identifiable predictive factors like age, schooling, years of experience, etc., may, of course, be utilized in so far as their predictive significance has been statistically demonstrated—but the use of definite predictive factors like these is not part of the interview. When specific observations from the interview can be shown to have similar predictive validity, they would then become usable and should be recorded.

The elements which enter into our estimate of a given quality are infinitely varied and numerous. There is an enormous range of functionally equivalent indicators of given traits and abilities. Significant comparison of candidates is not feasible by a simple direct comparison of the specific evidence, but should be feasible with respect to the traits inferred from the varied fragments of evidence.

If we place the emphasis on trying to satisfy legal requirements for review, I fear we shall merely induce interviewers to record elaborate observations which serve as convenient excuses for their ratings—and worse, we may cause the interviewers to shy away from some of their conclusions concerning a candidate simply because they are unable conveniently to specify items of evidence which led to the final evaluation.

All of this is sound in so far as it applies to patterns of behavior and traits that can be inferred from such isolated bits of evidence with predictive validity. The record can and should, in such a case, show that the ratings of the traits in question were based on varied evidences duly recorded and arrived at through a consideration of the interdependence of isolated items and their consistency with other recorded data about the individual. Such a record is not only possible, it is desirable; but it is possi-

ble only when the examiners are competent to recognize and evaluate that kind of data. No competent examiner should be afraid to state the basis of his rating on the record if it is statistically justifiable and valid. If he is not competent, he should not be allowed to conduct an examination of that type, however useful the technique is in competent hands. The pretense of maintaining legalistic objectivity goes hand in hand with the pretense of being able to "judge" personality subjectively; a competent record goes hand in hand with competent examining.

## *Chapter IX*

### **Appeal Procedure**

**I**N ANY process for the competitive selection of applicants for public employment, competitors should have full opportunity to observe the basis upon which the ratings were awarded and to appeal from the final determination when they believe that error has occurred. In some jurisdictions, an orderly method of appeal is provided from ratings on written tests, training, experience, and other parts of the examination, but not from oral ratings. However, the same reasons that demand a method for reviewing ratings on other parts of the examination also prevail in the case of the interview.

The principal reason why no appeal from oral test ratings has been provided in many jurisdictions is because there has not been preserved an adequate and reviewable record of the basis of the ratings, or of the method of application of the rating standards, if any, to the evidences of qualification. Such jurisdictions have logically felt that provision for appeals under the circumstances would encourage futile argument; for there can be presented no proof of error in the application of judgment to evidence that disappeared upon conclusion of the interview. But to deny review does not build confidence in any system. Thus public jurisdictions in the United States have come more and more to preserve a record of observations made during the oral test, of the questions asked and answers given, and of the basis of application of some rating standard to the evidence.

The mere provision that applicants may discuss their oral rating with the interviewers is insufficient to meet the requirements of a fair competitive process. There is an increasing demand for orderly procedure defined in rules or regulations and available to all applicants, whereby the applicant may look into the basis of the rating and present his claim of error as an appeal. This procedure should include a stated time limit within

which the appeal must be filed, a stated method of filing the appeal and method for its hearing; it should also provide adequate remedies for errors. When provision is made for an appeal within the examining agency, the machinery established for review must ordinarily first be used and the remedy therein provided exhausted before appeal may be taken to the courts.

### HEARING OF APPEALS

There are various methods of providing an internal review. The most elementary method used is to permit the applicant to appear in person before the interview board which examined him to discuss his case and assert his grievance. This requires considerable further expenditure of time by the interview board and seriously disrupts the work schedule of the board if it is a professional body carrying on continuing work. It is equally unfortunate to have to ask outside experts to give further time, as appeals come in, to haggle over ratings in a test which they believed had been completed. Moreover, appeals to the same individuals who rendered the disputed decision are seldom satisfactory to appellants. However, there are a number of additional possible channels of appeal, each of which will be discussed briefly.

#### *The Chief of the Examining Division*

Some of the larger jurisdictions provide that appeals from interview ratings shall be heard by the chief of the examining division, who has the primary and ultimate responsibility for the proper conduct of all examinations, and who is assumed to be familiar with everything that occurs in each test, from the selection of factors and standards to the method of their application. The chief of the examining division is not likely to feel that the appellant is criticizing him—a feeling which is believed to interfere with the administration of justice when the original judges are allowed to sit as reviewing officers on an appeal from their own decision. Again, however, the element of time required of a busy chief of the examining division, who is responsible for all elements of the tests, is likely to make this

solution of the appeals problem inadvisable. Moreover, he often feels the necessity of affirming the decision of his staff members.

### *The Commission or Administrator*

In some jurisdictions with a commission form of administration, the commissioners themselves sit as a board of appeals. The duties of a commissioner, however, are generally predominantly administrative, and administration in itself requires more time than most commissioners are able or willing to give. The judicial functions of a commission have been generally limited to interpretation of the law and rules in their administrative application. But in those jurisdictions where the time of commissioners can be made available, this method of internal review has been found more successful than the two previously described. In jurisdictions with a single administrator the time this official could devote to appellate hearings would be even more limited. There is serious doubt, too, whether "lay" commissioners will always have the understanding of requirements and standards on which to base an effective review.

### *The Board of Appeals*

Another solution of the problem is the establishment within the civil service agency of a board of appeals, independent of the examining division. In larger jurisdictions a board of appeals has proved to be successful in handling the grievances of applicants in every form of test; such a trained board is competent to consider on its merits each appeal from an interview rating. The chief objection to the board of appeals is that it takes some measure of responsibility from the examining division and may thus decrease the dependability of examining work. If the board of appeals, in turn, is to maintain its independence from the examining division, it should have investigators and examiners of its own who are competent in the field of its work.

### *The Courts or an Independent Tribunal*

When the machinery set up within a civil service agency has been used without bringing satisfaction to the applicant, the

next step (when fair and open competition is guaranteed by law) is to appeal to the courts or to some independent tribunal, judicial in its nature, set up for the purpose outside the civil service agency. There is some difference of opinion as to the need for outside review of the administrative work of a permanent agency such as a civil service commission. But strong argument can be made that such review is essential to maintain the confidence of applicants and prevent administration from becoming an arbitrary or routine function.

### TYPES OF ERROR

Many different grounds of error may be alleged by disgruntled candidates against oral examiners. Only the most common alleged causes of error will be discussed here.

#### *Evidence Overlooked*

The most frequent allegation is that the evidence presented, or that which should have been apparent, was overlooked in the rating process. This matter is one of fact and should appear in the record and be subject to objective review upon appeal. However, appellants sometimes contend that additional evidence of their qualification exists but was not called for at the time of the interview and that the applicants were disqualified or rated down because of the absence of evidence that they could have presented to prove their qualification. The proper answer to this argument on appeal from oral ratings, as in an appeal from all other ratings, is that the applicant has had his day in court. The burden of proof is always on the applicant, and if he has failed to present evidence of qualification at the proper time he must then hold his peace and wait for another examination. To reopen examinations or re-rate on the ground of newly discovered evidence or evidence not before presented is to prolong the competition unduly; this prolongation is unfair to other competitors. A claim of insufficient time allowed to present all of a candidate's evidence is an admission of failure to use the time provided as advantageously as other candidates. All candidates must be treated alike.

### *Mistake in Rating*

Another frequently alleged cause of error is that there has been a manifest mistake in the application of the rating to the evidence. When the rating scale is part of the record, when the evidence is preserved and the standard of application has been uniform, it is possible to discover this mistake by reviewing. Mistakes in rating do happen, although they are probably not common. Any valid system of appeal provides an opportunity to correct this kind of error, but correction should be allowed only when the alleged error is apparent on the record.<sup>1</sup>

### *Disclosed Identities*

Common allegation of disgruntled competitors is that one of the interviewers knew personally one or more of the candidates, and that the disclosure of the identity of certain candidates in the oral interview resulted in prejudiced rating. Many jurisdictions have a provision in their regulations that an interviewing examiner who personally knows one of the applicants shall not participate in the rating of the applicant. It has frequently been observed that when the examiner knows the applicant personally, he is more likely than not to lean over backwards in an effort to be completely objective; this practice, however, is as unfair as a prejudiced rating. It sometimes happens that in a test for a highly technical position the experts employed on the interview board know almost all of the applicants, and in this type of test it is often impossible to avoid the danger of error inherent in disclosed identity. For the general run of positions, however, proof that one of the examiners who rated knew a particular candidate personally should be a reason to scrutinize thoroughly the record of questioning, the evidence, and the rating awarded to ascertain whether any conscious or unconscious bias is evident which, if material, would require a re-examination.

<sup>1</sup> W. A. Hannig properly points out that caution is essential in correcting "mistakes" at a date some time after ratings were awarded. Application of a different standard of rating may result in a greater "mistake" than that the examiners sought to correct.

*Lack of Standards*

In jurisdictions where the judgment of examiners is the principal criterion of rating, disgruntled candidates will almost invariably allege that the oral test was not justly competitive in that there was no proper rating of factors and no relevant evidence adduced. Such an allegation, when substantiated by the record, has been ground for reversal in the courts.

*Irrelevant Objectives*

Another ground of error sometimes alleged is that the examiners have rated ideologies or irrelevant objectives prejudicial to some of the applicants. Almost all jurisdictions prohibit by law the consideration of the race, creed, or politics of applicants, and when the record on review discloses reasonable grounds to believe that these factors influenced rating, a new examination should be ordered. However, it is important to distinguish here between "politics" in the sense of adherence to a purely political party and the deeper implications of adherence to an ideology that may be prejudicial to the peace and safety of the community. Also, it is important to remember that a particular social outlook may be a ratable factor in determining qualifications for certain types of positions involving the exercise of broad administrative powers.

*Favoritism*

There are many allegations made in appealing which cannot be substantiated by the record. These allegations usually relate to vague charges of favoritism and prejudice—in short, that the examiners were influenced by considerations other than the evidence. It may be claimed that the examination was "in the bag" for a particular applicant. On appeal, of course, as in the presentation of qualifications, the burden of proof is always on the applicant. That fact is ordinarily enough to dispose of the appeal, but the best method of disproving such allegations is a full disclosure of the factors rated, the standards of rating used, and the actual application of those standards to relevant evi-



dence of possession of the factors adduced in the test and preserved in the record.

### REMEDY FOR ERROR

The object of appeal procedure is to provide a remedy for any possible error. Appeals from error are a proper protection for candidates, but it should be made clear that appeals do not ordinarily warrant a review of the honest exercise of discretion. Possible differences in judging the weight of evidence are never a proper ground for appeal; abuse of discretion must be clear and willful to support relief sought on that ground.

The discovery of reversible error in any examination raises difficult problems of administration. Frequently the register resulting from the test has been promulgated and appointments made before the appeal can be disposed of. The necessity for changed ratings after a register has been established causes untold grief to the civil service agency; yet the integrity and continuance of any civil service system depends upon its fairness and its readiness to admit and correct any error that has occurred, regardless of cost and difficulty.

#### *Re-Rating Applicant*

When the error relates to a mistake in rating a single applicant and when the evidence is sufficient to enable the board of review to ascertain the correct rating, the proper remedy is to re-rate the appellant and, if necessary, to change the rank order of eligibles. When the correct rating cannot be ascertained but an error has been established, some jurisdictions do and some do not authorize the retesting of a single candidate. A court will generally order a retest in spite of the difficulty of recreating comparable conditions.

#### *Re-Rating All Candidates*

When an error relates to the integrity of the rating standard, or the method of applying it to all the evidence, it may be necessary to re-rate all candidates and re-establish the register. This is rarely necessary if the examination was properly planned.

*New Examination*

When the error relates to the whole examination and, because of its gravity, requires the re-rating not only of the appellant, but of other applicants as well, it is possible to reorder the whole examination and reopen applications, but this method is seldom followed unless the condition of the labor market makes it advisable to start afresh and time is not pressing.

## SUMMARY

A proper method for appeal from oral test ratings is essential to the integrity of the testing process. The cost of providing an adequate review procedure is part of the necessary cost of civil service administration. The fact that errors are rare is not an excuse for avoiding the consequences of the few errors that occur. The fact that there is confusion, suspicion, and grief attendant upon any change in rating or the ordering of a new test or whole examination should not provide an excuse. Confidence in the system is of primary importance. Fairness is essential. An honest appeal procedure is the hallmark of good administration.

## Chapter X

### Legal Sanctions and Limitations

THE function of the court, in making an authorized review of any competitive examination, is to determine whether or not there has been formal and substantial compliance with constitutional or statutory provisions. Reported decisions in which the practice of civil service agencies with respect to oral testing has been reviewed are limited to a few jurisdictions. These decisions have taken a strictly legalistic position on the principles affecting validity in law under particular statutes and have sought to avoid passing upon the merits of techniques or devices employed. The following summary of court decisions seeks to collate existing judicial opinion, but the difference between a court's limited attention to validity in law and the extrajudicial requirement of test validity, which remains the additional problem of examiners, should be kept constantly in mind.

#### SCOPE OF THE DISCUSSION

Eliminated from consideration in this discussion are cases dealing with administrative defects common to all competitive examinations, such as fraud, mistake, and clerical error. These are governed by well-known principles of law, and are subject to common legal remedies regardless of the form of the test in which they occur. This discussion is concerned with the search of the courts for validity in law. The question has been frequently raised in the state of New York, where the courts have been seeking a formula by which the conduct of competitive oral tests can be judged legally valid or invalid. In this search, the courts have proceeded from a premise of principle: that the goal of the competitive process is to secure an *objective* test result. In seeking a standard of objectivity the courts quickly discovered that the competitive interview, or oral test, presents spe-

cific difficulties because it frequently seeks to measure intangible qualities which themselves defy exact definition. In measuring intangibles complete objectivity can rarely, if ever, be achieved, but the subjective element can be materially reduced. This possibility is recognized by the courts in the few cases where the question has been squarely presented. To prevent the test from representing largely the uncontrolled judgment of the examiner the New York courts now insist that the test employ some "objective standard or measure."<sup>1</sup> Having established this primary requirement of legal validity, the courts have been faced with the practical problems of application. What constitutes an objective standard or measure? To what extent must it be employed? It is with these questions that this summary is primarily concerned.

#### POWER OF THE COURT ON REVIEW

The formulation and conduct of the competitive oral test involves a series of determinations by a civil service commission or examining board. When the validity of the test is challenged, one or more of these determinations are reviewed by the court. The general rule that courts will not interfere with the determinations of administrative boards in the absence of allegations of bad faith or illegal action applies to these determinations.<sup>2</sup>

Thus, in the examination of applicants for public employment, the commission has a wide discretion with regard to the manner of performing its duties and exercising its powers. Unless the action of the commission is clearly illegal the court will not interfere.<sup>3</sup> An appeal may be addressed to the courts only when it is made to appear that the examiners have acted arbitrarily or illegally in the course of formulating or conducting the test.<sup>4</sup>

<sup>1</sup> *Matter of Sloat v. Board of Examiners*, 247 N. Y. 367 (1937).

<sup>2</sup> *Matter of Maxwell v. Civil Service Commission*, 169 Cal. 339, 146 Pac. 871 (1915); *Matter of Simon v. Maguire*, 204 N. Y. 253 (1912); *Matter of Allaire v. Knox*, 62 App. Div. 29, 168 N. Y. 642 (1901).

<sup>3</sup> *Pratt v. Rosenthal*, 181 Cal. 158, 183 Pac. 542 (1919). See also John F. Dillon, *Commentaries on the Law of Municipal Corporations* (5th ed. rev.), pp. 700-01.

<sup>4</sup> *Matter of Sloat v. Board of Examiners*, *supra*.

## STATUTORY PROVISIONS

Civil service statutes generally make no specific reference to the content and conduct of oral tests. However, civil service laws quite often specify that the examination be "practical in character" and that it relate to those matters which will fairly test the relative capacity and fitness of the persons examined.<sup>5</sup> In a few jurisdictions civil service laws refer specifically to oral tests. In Maryland, for example, "examinations held by commissions to establish lists of persons eligible for appointment will be written or oral or in the form of a demonstration of skill or any combination of these. . . ."<sup>6</sup> The absence of specific statutory provision for the oral test does not preclude its use. If a commission determines that an oral test is a "practical" means of testing knowledge or pertinent qualities the court will not interfere unless there has been a clear abuse of discretion.<sup>7</sup>

The federal courts have not as yet had occasion to rule upon the validity of the oral test as a part of examinations conducted by the United States Civil Service Commission. There is, of course, no federal constitutional provision concerning civil service examinations and the Commission operates under the statutory injunction stipulating that examinations "shall be practical in their character, and so far as may be shall relate to those matters which fairly test the relative capacity and fitness of the persons examined to discharge the duties of the service into which they seek to be appointed." In similar situations, the courts have ruled that the kinds of examinations employed by the Commission will not be questioned and that the exercise of the broad discretion vested in the Commission will not be challenged unless arbitrary, capricious, or unreasonable action is established by substantial evidence. To date, no one has

<sup>5</sup> This general requirement prevails in the state personnel statutes in Illinois, Maryland, Massachusetts, New York, and Ohio, and also in a number of local laws.

<sup>6</sup> Similar provisions for the oral test are found in the civil service laws of California, Connecticut, Rhode Island, Wisconsin, and other states.

<sup>7</sup> *Matter of Fink v. Finegan*, 270 N. Y. 356 (1936); *Matter of Sloat v. Board of Examiners*, 247 N. Y. 367 (1937). See also *Gollins v. Magee*, 121 N.J.L. 491, 3A, 2d, 581 (1939); *Junes v. Orlando*, 119 N.J.L. 227, 95A, 717 (1937).

alleged in the courts that oral tests conducted by the Commission have thus involved an abuse of the discretion vested in it by the civil service law and rules.

#### THE REQUIREMENT OF AN OBJECTIVE STANDARD OR MEASURE

The courts have reviewed competitive examinations in a number of cases in which it has been alleged that fraud, mistake, or other defect and error in administration has led to an invalid test result. It appears, however, that only in New York state has the competitive nature, and therefore the validity, of an oral test, as such, been challenged. Here the allegation in effect is that, aside from any question as to the validity of the test result, the examination as formulated and conducted precludes competition within the meaning of the governing constitutional and statutory provisions.<sup>8</sup> In these cases a court for the first time laid down a general rule which requires that in the measuring process an objective standard or measure be employed to eliminate in some degree the subjective element in the test result. The cases that have so far come before the court have involved oral tests, but the general principles laid down apply equally to all types of competitive tests.

In seeking to determine what constitutes an objective standard or measure it is essential to consider the rule laid down in the *Fink* decision which provides that:

An examination cannot be classed as competitive unless it conforms to standards or measures which are sufficiently objective to be capable of being challenged and reviewed, when necessary, by other examiners of equal ability and experience.

The import of this language is not entirely clear and its application has led to no little confusion. It may be helpful, however, in determining the scope and application of the rule, to

<sup>8</sup> Three cases have reached the New York Court of Appeals: *Matter of Fink v. Finegan*, 270 N. Y. 356 (1936); *Matter of Sloat v. Board of Examiners*, 277 N. Y. 367 (1937); and *Matter of Snyder v. Finegan* (Kern), 278 N. Y. 665 (1938). One case has come before the New York Appellate Division: *Matter of Bridgman v. Kern*, 257 App. Div. 420 (1939), affirmed 282 N. Y. 375 (1940). A number of cases involving examinations for teachers have been before the New York State Commissioner of Education. These cases follow without discussion the Court of Appeals' decision and, with one exception, will not be considered.

consider it in the light of test construction fundamentals.

It has been stated that the goal in any competitive examination is to secure an objective test result. This requires that a control of variables be maintained. Variables are interrelated factors which may adversely affect the uniformity of opportunity among the candidates and thus prevent the examiner from arriving at a valid and reliable rating of performance. The test must afford all candidates an equal opportunity to demonstrate fitness, and it must provide a yardstick which, if applied, will protect the rating against any arbitrary opinion of the examiner. *If the ratings are to be comparable, the test conditions and the test material must be reasonably invariable or constant as between candidates. Reliable evidence must be adduced on revealing and pertinent qualities to which must be applied a uniform standard of rating.*

It is obvious that to obtain an objective test result the measuring process must extend to each major determination in the formulation and conduct of the examination. The limits of permissible variation in the formulation and application of this standard or measure are not entirely clear. However, an examination of the decisions, in the light of what the court is seeking to accomplish, will indicate the framework, as outlined by the New York courts at least, within which a competitive oral test must be formulated and conducted.

### *The Application of the Objective Standard or Measure*

To obtain an objective test result the examiner must know what he is looking for. A mere statement of the duties to be performed will not in itself indicate the necessary qualifications, particularly when fitness depends to any extent on the possession of personal attributes. Thus two considerations are involved: (1) that qualities pertinent to the duties be specified, and (2) that there be some definition and understanding of what is included in the qualities selected. Only then can evidence of fitness be adduced from which comparable ratings can be made. The application of an objective standard or measure therefore requires a civil service commission or examining

board to make these two determinations preliminary to conducting the test.

### *The Determination of Qualities to be Tested*

The qualities to be tested must be important to the performance of the stated duties and the record on review must reveal a finding to this effect.<sup>9</sup> In criticizing the test reviewed in the *Fink* case the New York Court of Appeals said:<sup>10</sup>

The reason given for his (the candidate's) failure was that in the opinion of the examiners, although the candidate "was pleasant in manner and bearing and in comprehension fairly quick, he lacked force and executive ability and was altogether too mild." The record did not disclose that such qualities were important factors in the fitness of a physician in the positions for which the examination was held.

It is sometimes difficult to draw the line between an eligibility requirement and a qualification to perform stated duties. In either case, however, the courts will not interfere with the determination unless there has been a clear abuse of discretion.<sup>11</sup> However, in a recent New York case it has been suggested that a "point of view" on "political or quasi-political ideals" would be an improper qualification and would be excluded under the provisions of the New York state civil service law.<sup>12</sup>

The difficulties inherent in breaking down or defining qualities are recognized by the courts. From the language of the decisions it is often impossible to draw the line between the court's use of the terms "qualities" on the one hand, and "fac-

<sup>9</sup> *Fink v. Finegan*.

<sup>10</sup> *Sloat v. Board of Examiners*.

<sup>11</sup> *Pratt v. Rosenthal*, 181 Cal. 158, 183 Pac. 542 (1919); *Maxwell v. Civil Service Commission*, 169 Cal. 339, 146 Pac. 871 (1915); *Gunther v. Rice*, 246 App. Div. 179, 285 N.Y. S. 16 (1936). See also Dillon on Municipal Corporations, *op. cit.*, pp. 700-01.

<sup>12</sup> *Matter of Bridgman v. Kern*, 257 App. Div. 420 (1939), affirmed on other grounds 282 N.Y. 375 (1940). In this case the record did not disclose a finding that a "liberal" point of view had been predetermined as a qualification. However, the court indicates that under Section 25 of the civil service law "there should be no inquiries made by the examiners into the political and quasi-political ideal of candidates."

*Note:* It is important to distinguish between a prohibited consideration of partisan political ideology and a proper consideration of social objectives relevant to the administration of the duties of some jobs.—Eds.



tors" on the other. What the courts sometimes call factors are frequently as elusive as the qualities. Moreover, it is not clear under what circumstances or to what extent the court will require a breakdown or definition. In reading the cases it is well to keep in mind that the purpose of the court is to require the examiners, in so far as is practicable, to have a clear understanding of what they are looking for in the candidate's performance.

The foregoing considerations warrant a quotation from the court's opinion in the *Sloat* case. Commenting on "force" and "executive ability," the factors or qualities involved in the prior *Fink* case, the court said:<sup>13</sup>

It may be doubtful whether the nature or the element of these complex qualities can be satisfactorily defined, and, in saying that the candidate was, in the opinion of the examiners lacking in "force" and "executive ability," the examiners disclosed only their unsupported conclusion that the examinees lacked imponderable and undefined qualities without which the candidate was not fit for the position for which he was competing. . . . Exact definition of the qualities which are essential or desirable may be impossible; exact formula or standard by which such qualities can be measured has never been achieved; mechanical application of any standard is certainly not practicable.

The *Sloat* opinion dispels any disposition to regard the *Fink* case as requiring a noncompetitive test of qualities which cannot be satisfactorily defined. The character of the test remains discretionary with the commission or examining board and "much must be left here to the judgment of the examiners."<sup>14</sup>

### *The Notice or Announcement*

It has been stated that the application of an objective standard or measure breaks down unless all candidates have an equal opportunity to demonstrate their fitness. Candidates should therefore have some knowledge of the character and scope of the examination if their performance is to be a fair demonstration of their fitness. For example, it might be a serious handicap to certain candidates if they did not know whether they were to

<sup>13</sup> *Sloat v. Board of Examiners.*

<sup>14</sup> *Ibid.*

undergo a written or an oral test on a given quality. Generally, civil service laws and rules of civil service commissions make provisions for what must be included in the notice.<sup>15</sup>

Under the decisions it is clear that the character of each test and the qualities to be tested must be included in the notice of examination. In construing the broad provision of the New York civil service law and the rules of the New York Municipal Civil Service Commission, the Court of Appeals in the *Fink* case required the publication of "preliminary requirements and subjects" of the examination. Moreover, it was held that if any part of the examination was to be noncompetitive this fact likewise had to be announced.

The question arises as to whether the definitions or factors of the qualities to be tested must be included in the notice.<sup>16</sup> In the *Sloat* case it was claimed that the petitioner was prejudiced because of the insufficiency of the announcement. Candidates were merely advised that there would be a test in "oral English." Neither the factors thereunder nor the weights assigned to the various parts of the test were announced. The Court of Appeals referred this matter to the Commissioner of Education for final determination, and the examining board was upheld.<sup>17</sup>

<sup>15</sup> New York Civil Service Law, Section 14. See the discussion of the "Notice" under San Francisco Charter Provision, in *Pratt v. Rosenthal*, *supra*.

<sup>16</sup> In the *Fink* case the court, in discussing the factors of personality, said: "Unless the Commission can show that the test of force and executive ability was objective or that it properly exercised its discretion and determined that a non-competitive test is necessary and gave notice thereof, the examination of the petitioner should be re-rated." It would seem that this reference to notice refers only to the determination of a noncompetitive test and does not require the inclusion of the factors or definition.

<sup>17</sup> *Appeal of Florence M. Sloat*, 57 N.Y. St. Dept. Rep. 426, 429. "If the Board of Examiners had attempted to examine this candidate in fields not set forth in the announcement, perhaps the candidate would have some cause for complaint. There is no claim that this has been done. I find no necessity for the Board of Examiners to explain in its announcement of the examination in meticulous detail the particular subject matter to be covered, nor do I see any necessity of outlining to candidates the relative weight to be given any portion of an examination and the method of marking in connection therewith. The Board of Examiners is well within the rights under the statute in establishing a uniform system of marking and evaluation of the various items of an examination and applying the same equally to all candidates in the same examination. In any event, there is no legal requirement that the Board of Examiners provide this information in its announcement and its failure to do so is quite insufficient as a ground for setting aside the examination."

*Test Materials*

The materials of any test are the tools by which the evidence is adduced. In competitive examinations, test materials used in various parts of the examination usually take the form of written or oral questions. It is fundamental that questions in any competitive test must be reasonably calculated to develop evidence indicative of a candidate's fitness to perform the stated duties.<sup>18</sup>

It has been shown that qualities, in so far as is practicable, must be defined or broken down into factors so that the examiner may know more accurately what he is seeking. If the process of breaking down qualities into factors is to serve any function, the questions or test materials should be formulated with the view of adducing evidence related to the factors. The extent to which relevant evidence can be adduced as to the factors determines the degree of objectivity possible in the test result. This is borne out in the *Fink* decision in which the court held the test defective because it did not appear that the questioning was directed to the factors of personality to be tested,<sup>19</sup> and in the *Bridgman* decision in which the Court of Appeals stated: "Commentaries elicited from candidates were so vague or remote in character and broad in scope that the ratings in respect of 'soundness of position taken' disclosed only the unsupported conclusions of examiners."

It does not follow, however, that relevant evidence can be adduced only by questioning or that every question must be related to the factors or qualities tested. In testing qualities which can be evidenced only by observation, the test materials may be used as a means of observing these qualities provided the record reveals the basis of the rating.<sup>20</sup>

<sup>18</sup> Early cases required that questions be "relevant" to duties. *Darling v. McGuire*, 70 Misc. 597, 129 N.Y.S. 385 (1911).

<sup>19</sup> See the discussion in *Sloat v. Board of Examiners*, pp. 371, 373.

<sup>20</sup> For example, in *Sloat v. Board of Examiners*, the personal attributes of candidates were tested by questions related, not to the factor being tested, but to other test materials consisting of reading selections. However, both answers and observations, which formed the basis of the rating on personal attributes, were recorded in the test record.

The New York courts do not insist, in the adduction of evidence, that all candidates be asked all of the same questions in the same form. The important consideration is that the questions asked were reasonably calculated to adduce comparable evidence from which comparable ratings might be made. An examination of the *Sloat* and *Snyder* decisions shows that the court will allow a considerable degree of discretion in the application of test materials. In the *Snyder* case each candidate was asked the same key questions, but the supplementary or follow-up questions depended upon the answers to the key questions and, therefore, differed as to each candidate. The court, on the final order directing a verdict for the civil service commission, determined that reasonable equality as between candidates had been preserved in that the questions were "similar."<sup>21</sup>

In the *Sloat* case, on the other hand, the test material differed as between candidates. For example, in the oral English test all candidates were not given the same reading selections nor were they asked to pronounce the same lists of words. In sustaining the oral English test the court said:<sup>22</sup>

Evidently it is not practicable to apply such tests in exactly the same form to each competitor or to make exact comparisons between them. This is true in some degree of every examination, especially of examinations calculated to show intellectual ability and broad cultural learning.

While these decisions allow the examining board considerable discretion, a word of caution is warranted. If the test materials differ as between candidates there should be a *determination* as to their comparable difficulty. In the *Sloat* case the examining board made a determination that the various rating selections were of equal difficulty and reasonably calculated to adduce evidence from which comparable ratings could be made. The board would be well advised to record the determination in its minutes.

<sup>21</sup> Mr. Justice Valente, Special Term, Part III, *New York Law Journal*, December 26, 1936, Vol. 96, No. 147.

<sup>22</sup> *Sloat v. Board of Examiners*, p. 373.

### *The Standard of Rating*

It is commonly understood that a valid test can be secured only if a uniform standard of rating is applied to the evidence adduced. No case has been found which specifically discusses the essential requirements of a valid rating scale, though in a recent case a New York court questioned the test result on a basis of lack of evidence of existence of a uniform standard applicable to the evidence adduced.<sup>23</sup> In the *Snyder* case the trial court avoided the issue in the following language:<sup>24</sup>

While it did not appear that the examiners representing the Commission had adopted scientific and mathematical standards in weighing elements entering into personality, I did not think the proof was sufficient to authorize me to say that a particular set of mental measurements which can be reduced to some form of mathematical accuracy exists to such an extent that the court could take notice of it as a scientific principle generally recognized. The issue, in my opinion, resolves itself into whether the examiners had treated all the candidates equally, without bias and prejudice, and whether they had adopted a similar set of questions for each person examined.

This statement probably represents the limit to which any court will go at this time. It may look to see if the examiner "treated all candidates equally" but it will not require the application of a specific rating scale. If from the evidence adduced the evaluation of a given factor or quality seems reasonably fair the court will not interfere.<sup>25</sup>

### THE RECORD FOR REVIEW

It will be recalled that the rule laid down in the *Fink* case requires that the competitive test conform to measures or standards which are sufficiently objective "to be capable of being challenged and reviewed, when necessary, by other examiners of equal ability and experience." It will be observed that this

<sup>23</sup> *Bridgman v. Kern*, *supra*.

<sup>24</sup> *Matter of Snyder v. Finegan*, *supra*.

<sup>25</sup> In *Sloat v. Board of Examiners* the court concludes that the mechanical application of any standard is not practicable and that the "result may depend as much upon the fitness of the examiners as upon the fitness of the candidate. That is a risk inherent in all systems of examination."

requirement does not extend to the merits of the test result. Conceivably, an objective standard or measure could be applied and a valid test result obtained without any test record being preserved. The court recognizes, however, that the test result rests largely on determinations and evidence peculiarly within the knowledge of the examiners. If an aggrieved candidate must base his challenge largely on the recollections and interpretations of the examiners, he has, in effect, lost his right to review. Moreover, the court of review must determine whether the test as formulated and conducted complies with constitutional or statutory provisions. Therefore, the court must have before it a test record which an independent examiner would need in order to determine the basis of the test result.

A number of questions arise as to the sufficiency of the record. In the main, however, they present no special problems. This summary will be limited to the one novel question which has so far been presented to the courts: What constitutes sufficient evidence of observations of candidates' performance?

Turning first to a comparison of the *Fink* and *Snyder* cases we find two main distinguishing features. In the first place, the questions asked the candidates in the *Fink* case were not related to the quality to be tested, that is, they were questions on surgery and general medicine and were not, in themselves, intended to elicit answers from which the examiners could judge the qualities they were testing. The questioning was merely a device by which the examiners could see, hear, and observe the candidate and judge his personality. The rating was based upon these observations. The candidate's performance with respect to the qualities tested was nowhere preserved. In the *Snyder* case, however, the questioning was related solely to the qualities tested. The questions were intended to elicit from the candidate answers from which the examiners could judge the candidate on a basis of his stated prior demonstration of possession of the factors involved. A stenographic record of the questions and answers was preserved. Thus the two cases were distinguishable; in the former there was no record of relevant evidence, and in the latter all the evidence rated was in the record.

The *Sloat* case, however, presents a different situation. In neither the oral test nor the teaching test was there a complete record of the candidate's performance upon which the examiners based their rating. As in the *Fink* case the test result was based upon what was seen, heard, and observed. However, the examiners in the *Sloat* case indicated, largely by brief notations on the rating sheet, their observations of the candidate's performance. These notes were preserved and were before the court on review. From an examination of the exhibits, however, it is doubtful whether an independent examiner could have determined with any degree of accuracy the basis of the rating.

The *Sloat* decision makes no specific reference to what constitutes sufficiency of the record of evidence rated. It merely finds a sufficient record in that it sustains the Board of Examiners and concludes that the tests as formulated and conducted were reasonably calculated to reveal the relative merit and fitness of the candidates. Therefore doubt remains as to what constitutes a sufficient record of evidence for cases in which the rating is based upon what the examiners see, hear, and observe. At this point it can only be said that in such cases an examiner should record, in so far as is practicable, a detailed resume of his observations so as to be able to show that the ratings were based on substantial evidence developed and not merely the arbitrary or capricious action of the examiners.

In conclusion, it should be remembered that courts in few states have given judicial consideration to the problems of oral examining; in New York, where such consideration is most frequent, the courts are still feeling their way toward a formula that will distinguish procedures "sufficiently objective" to afford reasonable opportunity for review from those which permit unsupported conclusions of examiners. The danger of requiring excessive formalism, in the name of "validity in law," has been impressed in recent years upon the courts; the impropriety of excessive subjectivity has been impressed upon examiners. The course of future practice and decisions should round out the complicated legal picture that exists today and point to a workable relationship between law and practice.

## *Chapter XI*

### **Evaluation of the Oral Process**

**S**PORADIC efforts to evaluate the competitive oral test process in a small number of jurisdictions have disclosed little evidence of the validity of the types of interview they use. However, this does not necessarily constitute a reflection on the worth of these processes. Very often the lack of evidence of validity springs from a related lack of objective yardsticks for determining success on the job—the final and most conclusive basis for appraising the value of any selection process. Until a jurisdiction has established a dependable system of service ratings that may be used as a basis for measuring the validity of its selection methods it must depend on indirect measures. Conclusions drawn from such studies are largely inferential, and this fact should be kept in mind when findings are appraised.

#### **CORRELATION WITH GRADES IN OTHER SUBJECTS**

The correlation of oral test ratings with grades assigned to the same candidates in other tests in the general process of measuring qualifications proves little about the effectiveness of any of the processes used. This device fails primarily because interviews are used mostly to measure aspects of qualification that do not lend themselves to other forms of testing. Moreover, the interview has been used too often as a means of eliminating individuals who, although capable of passing tests of skill or knowledge with a high score, still lack or are believed to lack essential characteristics for appointment.

In a recent study of an examination for civil service examining assistant, conducted by the New York Municipal Civil Service Commission, a correlation of  $+ .05$  was found between oral and written tests—a coefficient well within the limits of chance, and which may be taken as zero for practical purposes. The conclusion of the research examiner who conducted the



experiment was that the interview measured an ability other than the informational background and clerical capacity appraised in the written test. Correlations between experience ratings, based upon written statements of previous employment verified by vouchers and oral test ratings were found in the same study to be  $-.22$ , a coefficient which again may be taken for practical purposes as zero. The correlation of the oral test with written and experience tests was by rank coefficient.

Such studies cannot be conclusive evidence that the interview failed of its purpose; for it is possible also that the factors rated in the written and experience tests are themselves inferior measures of what they purport to test. Moreover, there is little reason to expect that such correlations will disclose much about the value of either type of test, because the interview is designed to rate wholly different evidence on wholly different factors of qualification. When the interview is designed to implement previous test ratings and to develop additional evidence of knowledge, training, or experience, the tests are more likely to have some basis of logical comparison; only a few jurisdictions have employed the oral test primarily to retest qualifications previously rated by other methods.

#### CORRELATION WITH PERFORMANCE ON THE JOB AND PROMOTION RECORDS

Efforts to evaluate oral tests by comparison with subsequent performance ratings or promotion records of employees have met with little success. Many of the systems for evaluating the service value of employees are based upon the judgment of supervisors unsupported by recorded evidence. Frequently the supervisors have no common standard for performance rating. As to promotion criteria, by the time employees have served sufficiently long to develop promotion records, the record on which oral test ratings were based at the time of the original appointment has too often been filed away or lost. Again, the basis for promotion in many jurisdictions is not uniform throughout the various governmental units, and frequently promotion is not based upon an objective or reviewable record.

Unless known bases for criteria exist and can be compared there is little practical value in correlations.

In the New York City study of the oral ratings in the examination for civil service examining assistant, a correlation of  $+.01$  between service ratings and the oral test was found.<sup>1</sup> This coefficient also may be taken for practical purposes as zero. The report of that study concluded with this statement: "If the service rating rates efficiency, the oral yielded no information enabling a forecast of candidate's efficiency."

### STUDIES OF VALIDITY OF JUDGMENT RATINGS

In 1936 an experiment in oral testing was conducted in England under conditions approximating those usually obtaining in the *viva voce* examination in England and in some American interview tests.<sup>2</sup> A parallel to the most usual form of American interview is evident in the objective to test "alertness, intelligence, and general outlook." The purpose of the experiment was to test the degree of consistency of two examining boards appointed to conduct an examination of this kind.

Subjects were obtained for this experiment by offering a substantial award to the best candidate in the examination. Applications were accepted from recent university graduates or from students certified by the university authorities under whom they had worked as being suitable candidates for the junior grade of the administrative class (Home Civil Service). Candidates were informed that on the basis of the interview the examiners would judge the value of the candidate's personality for the kind of career that the Home Civil Service offered. There were two boards of examiners. The raters were all supplied with the same confidential report regarding each candidate from a tutor or other authority, along with the candidate's own statement as to his background. Another form provided details regarding the candidate's schooling, employment, and the like.

<sup>1</sup> The correlation of the oral test with service rating was by use of the Eta coefficient.

<sup>2</sup> P. Hartog, and E. C. Rhodes, *The Marks of Examiners* (London, 1936). The present description of this experiment constitutes part of an unpublished review of *The Marks of Examiners* submitted to this committee by Dr. Norman Powell, formerly of the staff of the New York Municipal Civil Service Commission.

Candidates were examined for not less than a quarter of an hour and not more than half an hour. There were 16 candidates divided into two groups. Those in Group A appeared in alphabetical order before Board I and then before Board II. Those in Group B appeared first before Board II and then before Board I. Examiners set down their scores independently before discussion. After discussion, a final mark was recorded which reflected the opinion of the examiners as a whole. This mark was obtained either by agreement, or when this was impracticable, by averaging the ratings of the individual examiners.

The results of the experiments are striking. The coefficient of correlation between the ratings of the two boards was  $+ .41$ . It is stated that "The extreme difference between the two boards' estimates of the candidates' merits, amounting to 20 to 30 marks out of 100, and the average difference of about 12 marks out of 100, point to the unreliability of the interview test." One is impressed by the finding that "the actual evidence produced seems to have been so different that we might almost have supposed different candidates to have been examined." The average correlation between the marks of the individual examiners and the final award of their particular board was  $+ .83$  for the first board and  $+ .78$  for the second board. In a postscript, Hartog points out that the mode of approach of the two boards seemed to be identical, that the two boards were equally skillful in cross-examining, but that "it was largely a matter of chance whether they struck on a topic in which the candidate felt so strongly that he was able to display his individuality." In concluding his comments on this experiment, Dr. Powell states: "The field of oral examining remains a practically unworked terrain."

American studies of subjective interviews are not much more encouraging. The low reliabilities found for ordinary interview methods in industry are well known,<sup>3</sup> and a similar situation prevails in the academic field. According to one investigation, a graduate student was given four separate 50-minute oral tests on the content of a master's thesis; two committees passed

<sup>3</sup> M. S. Viteles, *Industrial Psychology* (New York, 1932), pp. 159-79.

the candidate and two failed her.<sup>4</sup> In another study, six graduate students were given a 45-minute oral examination on a college course in the psychology of adolescence.<sup>5</sup> The test was administered by three different examining committees, each consisting of four graduate students. The correlation among the committee ratings was  $\pm .30$ .

L. J. O'Rourke's effort toward standardizing oral examination procedure and his work in training examiners to make comparable ratings and accurate, complete, and uniform reports of the entire examination was an important first step toward more objective interviews.<sup>6</sup> O'Brien's work for the United States Civil Service Commission in requiring and analyzing probative evidence in oral tests should provide a basis for interesting new studies of validity.<sup>7</sup> In other jurisdictions many recent approaches to more objective oral testing and the development of standardized procedures have begun to build up a new tradition and method in this field; complete records are being preserved and more objective performance ratings developed for comparison.<sup>8</sup> Studies of the results of modern oral practice should reveal more positive findings in the next few years.

### THE PROBLEMS AHEAD

A review of the somewhat meager body of research thus far published on the reliability and validity of the oral test does not afford much ground for reaching any blanket conclusions for or against the interview as a tool of personnel selection. One fact to which the evidence does point, and on which this report has elaborated, is that the interview is quite susceptible to a number of influences that introduce or increase the factor of error in the over-all results. These influences are both internal

<sup>4</sup>S. L. Pressey, L. C. Pressey, and E. J. Barnes, "The Final Ordeal," *Journal of Higher Education*, April 1932, pp. 261-64.

<sup>5</sup>S. L. Pressey, and E. J. Barnes, "The Reliability and Validity of Oral Examinations," *School and Society*, November 23, 1929, pp. 719-22.

<sup>6</sup>Civil Service Commission, *Forty-fifth Annual Report*, 1928, pp. 37-49.

<sup>7</sup>J. C. O'Brien, formerly Principal Oral Examiner, United States Civil Service Commission.

<sup>8</sup>For an interesting analysis of the effect of the rating range in the oral test on final rank-order of candidates, see W. V. Bingham, *Oral Examinations in Civil Service Recruitment* (Civil Service Assembly, Pamphlet No. 13, 1939).

(halo, bias, lack of training, etc.) and external (lack of advance planning, misdirection of emphasis, etc.). Taken singly and collectively, they present a challenge that should certainly be met and overcome before the personnel examiner is ready to place full confidence in the oral test as a dependable tool of measurement.

### *Need for Rational Appraisal Standards*

There is also the related problem of reaching agreement on what constitutes the desirable characteristics of correlation between the oral test and other measures, in order that there can be set up a standard for determining whether the interview has performed its function in a satisfactory manner. The determination of such a standard awaits the completion of a good deal more research. In arriving at this standard, however, it would be well to keep in mind that the interview is often called upon to measure elusive and intangible values, and that interview groups are not always large enough to meet the requirements of a broad statistical sample. Because of these considerations, it may be asking too much of the oral test to expect micrometric precision of results with unfailing regularity.

Whether or not the oral test should be expected to show significant positive correlation with ratings on other parts of the over-all examination depends on the role assigned to it. If the oral test is looked upon as an additional means for appraising factors that are covered elsewhere in the examination, a fairly high degree of internal correlation should rightfully be expected. But this does not follow if those who plan the test intend that it concentrate on factors not measured in other parts of the examination. Here, a high positive internal correlation is evidence that the oral test has succeeded only in duplicating a function already performed elsewhere. Thus it is necessary that the researcher keep in mind the intended objectives of the oral test when he interprets correlation studies of this sort.

Regarding the validation of oral test ratings against an independent criterion of employee performance on the job, it is reasonable to look for a significant positive correlation between

ratings on factors shared in common by both measuring instruments.<sup>9</sup> But to regress a step further, there is still much to be done in the way of perfecting the independent criterion against which the results of the oral test can be appraised. At more than one point in the preceding pages the present-day shortcomings in the field of employee performance appraisal has been cited and lamented. Until those shortcomings are eliminated, any effort to determine the validity of the oral test by using an external criterion will be largely a process akin to that of lifting one's self by the bootstraps.

### *Need for a Realistic Attitude*

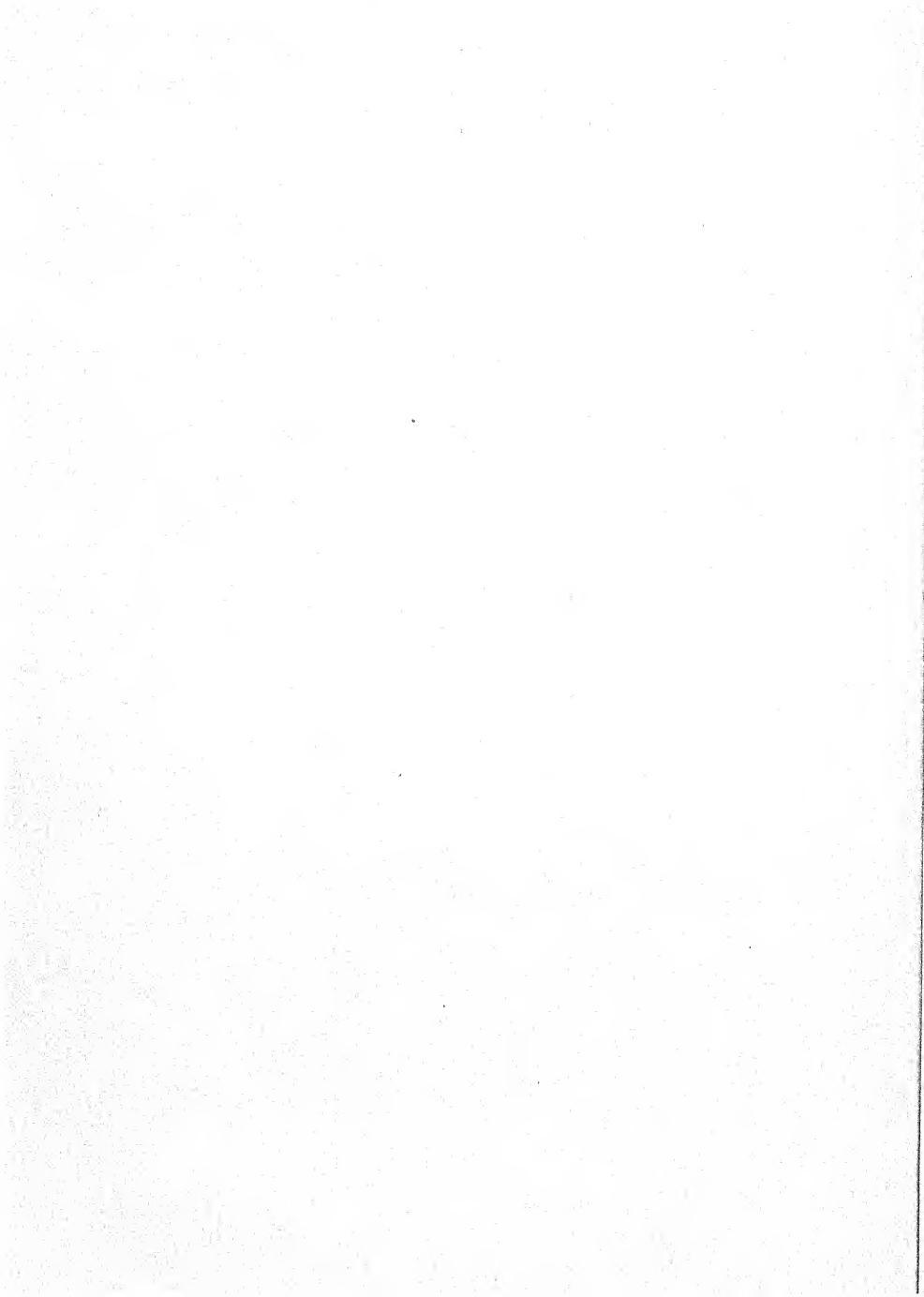
These are some of the obstacles that lie in the path of those who would seek to perfect the oral test as a step in the process of personnel selection. In recognizing them, this committee does not wish to imply that the oral test should remain unused until they have all been overcome. Such a view would be highly unrealistic, for the oral test will undoubtedly be brought to its highest eventual state of development by those practitioners who, while conscious of its present limitations, continue to use the interview with discretion and constantly try to improve it. While careful experimentation and research will continue to point the way toward better procedures, the translation of research into results can come only through the medium of actual use under operating conditions.

Until this evolutionary process yields fruit, it may often be necessary to answer the clamor of unfriendly critics for conclusive proof that the oral test does what it purports to do. Those who come to the defense of the oral test need not be dismayed

<sup>9</sup> In this connection it is quite possible that those who seek to validate oral test scores against over-all service rating scores overlook some important considerations. For example, if the oral test is designed only to evaluate certain factors of appearance and personality, then the proper validation criterion is not the over-all service rating score, but only that portion of it to which the same factors of appearance and personality contribute. In other words, it is unreasonable to expect close correlation between one measuring device, concerned only with appearance and personality, and another device which is concerned not only with appearance and personality factors, but also job knowledge, skill, industry, punctuality, and other qualities. Statistical integrity, it would seem, calls for the isolation of traits shared in common by the oral test and the service rating form as a first step in validating the former against the latter.

by the present dearth of positive evidence in its favor. While there is yet lacking a body of objective positive evidence that the oral test performs its allotted role invariably and unerringly, there is ample negative evidence to be cited—evidence of what can and does happen when the oral test is *not* included in examinations for certain types of positions. When the argument is placed on this basis there are few who will not concede that the oral test contributes constructively to the selection process.

Not all criticism of the oral test should be considered unfriendly, however, for much of it represents a sincere desire to call attention to deficiencies that are often glaring. Such criticism may well be taken in the constructive spirit in which it is offered. This report has recognized frankly that the oral test has in the past been abused as well as used. By pointing out shortcomings and by suggesting positive steps for overcoming them, the Committee has sought to bring enthusiast and critic closer to a common middle ground whereon both accept the oral test for what it is, and nourish it for what it will become.





## Appendix A

### Bibliography

- ALLPORT, G. W. *Personality: A Psychological Interpretation*. New York: Henry Holt & Co., 1937. 588 pp.
- BARNARD, CHESTER I. *Functions of the Executive*. Cambridge: Harvard University Press, 1938. 334 pp.
- BARNES, E. J. and PRESSEY, S. L. "The Reliability and Validity of Oral Examinations," *School and Society*, November 23, 1929, pp. 719-22.
- BINGHAM, W. V. "Halo Invalid and Valid," *Journal of Applied Psychology*, April 1939, pp. 221-28.
- BINGHAM, W. V. *Oral Examinations in Civil Service Recruitment*. Chicago: Civil Service Assembly, Pamphlet No. 13, 1939. 30 pp.
- BINGHAM, W. V. and FREYD, MAX. *Procedures in Employment Psychology*. New York: McGraw-Hill, 1926. 269 pp.
- BINGHAM, W. V. and MOORE, BRUCE V. *How to Interview*. New York: Harper and Brothers, third edition revised, 1941. 263 pp.
- FREYER, DOUGLAS. *The Measurement of Interests in Relation to Human Adjustment*. New York: Henry Holt & Co., 1931. 488 pp.
- HARTOG, SIR PHILIP JOSEPH and RHODES, E. C. *The Marks of Examiners*. New York: Macmillan, 1936. 344 pp.
- HOLLINGWORTH, H. L. *Judging Human Character*. New York: D. Appleton & Co., 1922, 268 pp.
- HOPKINS, JESS T. and Others. *The Emergence of a New Public Employment Service*. Rochester: Public Employment Center, 1935. 345 pp. Chapter 12.

- HORCHOW, REUBEN. *Machines in Civil Service Recruitment, with Special Reference to Experience in Ohio*. Chicago: Civil Service Assembly, Pamphlet No. 14, 1939. 43 pp.
- HULL, CLARK L. *Aptitude Testing*. Yonkers: World Book Co., 1928. 536 pp.
- KLEIN, ALICE CAMPBELL. *Civil Service in Public Welfare*. New York: Russell Sage Foundation, 1940. 444 pp. Chapter 10.
- KORNHAUSER, A. W. and KINGSBURY, F. A. *Psychological Tests in Business*. Chicago: University of Chicago Press, 1924. 194 pp.
- LAIRD, DONALD A. *The Psychology of Selecting Employees*. Third edition revised. New York: McGraw-Hill, 1937. 316 pp.
- LANE, MARIE DRESDEN. "The Education of an Examiner," *Survey Midmonthly*, April 1938, pp. 105-07.
- LINK, HENRY C. *Employment Psychology*. New York: Macmillan, 1919. 440 pp.
- MERIAM, LEWIS. *Civil Service Testing for Social Work Positions*. Chicago: Civil Service Assembly, Pamphlet No. 9, 1937. 6 pp.
- O'BRIEN, JAMES C. and MARENBERG, PHILIP P. *Your Federal Civil Service*. New York: Funk & Wagnalls, 1940. 501 pp. Chapter 9.
- ORDWAY, SAMUEL H. JR. and O'BRIEN, JAMES C. *An Approach to More Objective Oral Tests*. Washington: Society for Personnel Administration, Pamphlet No. 2, 1939. 31 pp.
- ORDWAY, SAMUEL H. JR. and LAFFAN, JOHN C. *Approaches to the Measurement and Reward of Effective Work of Individual Government Employees*. New York: National Municipal League (published as supplement to *National Municipal Review*, October 1935, pp. 557-601).
- O'ROURKE, L. J. "Oral Examinations," in United States Civil Service Commission, *Annual Report*, 1928, pp. 36-50.

- ROETHLISBERGER, F. J. and DICKSON, W. J. *Management and the Worker*. Cambridge: Harvard University Press, 1939. 615 pp.
- STUART-BUNNING, G. H. "The Use of the Interview in Recruitment and Promotion," *Public Administration* (London), October 1937, pp. 433-40.
- SYMONDS, PERCIVAL M. *Diagnosing Personality and Conduct*. New York: Century Co., 1931. 602 pp.
- SYMONDS, PERCIVAL M. "Factors Influencing Test Reliability," *Journal of Educational Psychology*, February 1928, pp. 73-87.
- TRAXLER, ARTHUR E. *Use of Tests and Rating Devices in the Appraisal of Personality*. New York: Educational Records Bureau, Bulletin No. 23, 1938. 80 pp.
- UNITED STATES SOCIAL SECURITY BOARD. *Manual of Merit System Administration*. Washington, 1938. Chapter 17.
- VITELES, M. S. *Industrial Psychology*. New York: W. W. Norton & Co., 1932, pp. 159-79.

## Supplementary References : 1948

FOLLOWING is a brief supplementary list of material dealing with personnel interviewing that has appeared since the original publication of *Oral Tests in Public Personnel Selection* in 1943. While this list makes no pretense at completeness of coverage, it should be useful to the reader who is interested in reviewing some of the more recent relevant material in this field.

- BARNARD, MARGARET W. and BRODY, WILLIAM. "A New Method of Selecting Health Officers-in-Training," *American Journal of Public Health*, June, 1947, pp. 715-20.
- FEARING, FRANKLIN and FEARING, FLORA MAY. "Factors in the Appraisal Interview Considered with Particular Reference to the Selection of Public Personnel," *Journal of Psychology*, May, 1942, pp. 131-53.

FREEMAN, G. L. "Using the Interview to Test Stability and Poise," *Public Personnel Review*, April, 1944, pp. 89-94.

MANDELL, MILTON M. "The Group Oral Performance Test," *Public Personnel Review*, October, 1946, pp. 209-12.

MORSE, MURIEL, and HAWTHORNE, JOSEPH W. "Some Notes on Oral Examinations," *Public Personnel Review*, January, 1946, pp. 15-18.

OTIS, JAY L. "Improvement of Employment Interviewing," *Journal of Consulting Psychology*, March-April, 1944, pp. 64-69.

POWELL, NORMAN J. "Relation of Order of Appearance to Mark in Oral Tests," *Journal of Applied Psychology*, February, 1943, pp. 104-108.

SHNEIDMAN, EDWIN S. "A Note on the Experimental Study of the Appraisal Interview," *Journal of Applied Psychology*, April, 1943, pp. 196-205.

UNSIGNED. "An Experiment with the Oral Examination," *The Compass*, February-March, 1941, pp. 6-8.

UNSIGNED. "A Good Man Is Hard to Find," *Fortune*, March, 1946, pp. 92-95, 217-18, 220-23.

## *Appendix B*

### **Rating Forms**

Examination Center		STATE OF ALABAMA—PERSONNEL DEPARTMENT		Identification Number		SCORES	
Date		ORAL INTERVIEW RATING FORM		Do Not Write in these spaces			
To Members of Examining Board: When the candidate has left the room rate him (or her) on each of the following points and place your rating in the space provided. Do not make any marks on the reverse side of this form.							
1. WOULD THE CANDIDATE'S APPEARANCE (NEATNESS, ETC.), CREATE A FAVORABLE IMPRESSION UPON PERSONS WITH WHOM HE WOULD DEAL?							
unfavorable impression, unattractive appearance	unimpressive, flabby or untidy	suitable, acceptable, neat and in good taste	creates distinctly favorable impression	impressive, commands admiration			
2. IS THE CANDIDATE'S VOICE AND ENUNCIATION SATISFACTORY FOR A POSITION WHICH FREQUENTLY INVOLVES CONTACT WITH OTHERS?							
sturred, indistinct or irritating	understandable but unpleasant	neither unpleasant nor unpleasant	definitely pleasant and distinct	exceptionally clear and pleasing			
3. IS THE CANDIDATE'S MANNER SUCH THAT HE EASILY ESTABLISHES FRIENDLY RELATIONS WITH THOSE WITH WHOM HE COMES INTO CONTACT?							
too aggressive, ill-mannered or too reserved	not very responsive, lacks initiative	friendly, polite and likable	very cooperative and well-mannered—considerate	inspires confidence; establishes pleasant relationship readily			
4. DOES THE CANDIDATE SHOW GOOD EMOTIONAL CONTROL? IS HE AT EASE IN THE PRESENCE OF OTHERS?							
too timid, adjusts to situation with difficulty	self-conscious, ill at ease, or lacks restraint	poor control, self-confident, no tension	well poised, stable, at ease	exceptional poise, adjusts to situation without effort			
5. IS THE CANDIDATE ALERT IN COMPREHENDING QUESTIONS AND SITUATIONS?							
often misunderstands questions, fails to grasp the situation	slow to grasp questions, requires explanations	usually grasps intent of questions, responds readily	grasps situation quickly, anticipates questions	exceptionally, has unusual perception			
6. IS THE CANDIDATE ABLE TO PRESENT HIS IDEAS COHERENTLY AND LOGICALLY?							
confused, illogical, naive or biased	involved, self-contradictory, introduces irrelevances	logical presentation, some evidence of judiciousness	weighs situations carefully, superior reasoning ability	evaluates situations critically, keenly and accurately			
PERSONAL REMARKS: In the field of the position described, how do you rate the candidate's suitability for the position? For what he is probably doing in mind the duties of the position, would you recommend him? Indicate what you believe to be an objective rating by placing a check in the appropriate space.							
POSITION:				Do Not Recommend		Recommend	
				With Hesitation		With Confidence	
				Recommend With Enthusiasm		Score (Leave blank)	
1.							
2.							
3.							
4.							
5.							
6.							
Is this candidate known to you?							

Form No. 51  
2-28-35-30-36

EXAMINER'S SIGNATURE

FIGURE 1. RATING FORM (Alabama State Personnel Department)

<b>COMMENTS:</b>	
1. APPEARANCE:	
2. VOICE:	
3. MANNER:	
4. EMOTIONAL CONTROL:	
5. ALERTNESS:	
6. ABILITY TO PRESENT IDEAS:	
TO WHAT EXTENT IS THE CANDIDATE KNOWN TO YOU?	

FIGURE 1. (Reverse)

# RATING FORM FOR USE OF INTERVIEWERS AND ORAL EXAMINERS—2—1938

**INSTRUCTIONS:** Ask yourself how this applicant compares with those who are doing work of this kind. Consider whether his voice, appearance, etc., would be a liability or an asset in such a position. Rate him by making a check (✓) at that point on each scale where, in your judgment, the applicant stands. Rate the following traits:

- 1. VOICE AND SPEECH.** Is the applicant's voice irritating, or pleasant? Can you easily hear him? Is his voice clear and distinct? Is his voice so harsh, raspy, and well-modulated that it would be a valuable asset in this position?
- 2. APPEARANCE.** What sort of first impression does he make? Does he look well-groomed and neat? Is he well-groomed or slovenly? Erect or slouchy? Attractive or unattractive in appearance?
- 3. ALERTNESS.** How readily does he grasp the meaning of a question? Is he slow to understand or does he understand quickly? Does he understand the meaning of the question? Is he slow to understand or does he understand quickly? Does he understand the meaning of the question?
- 4. ABILITY TO PRESENT IDEAS.** Does he speak logically and convincingly? Or does he tend to be vague, confused or illogical?
- 5. JUDGMENT.** Does he impress you as a person whose judgment would be dependable even under stress? Or is he hasty, erratic, biased, swayed by his feelings?
- 6. EMOTIONAL STABILITY.** How well poised is he emotionally? Is he touchy, sensitive to criticism, easily upset? Is he irritated or impatient when things go wrong? Or does he keep an even keel?
- 7. SELF-CONFIDENCE.** Does he seem to be uncertain of himself, hesitant, lacking in assurance, easily bluffed? Or is he wholesomely self-confident and assured?
- 8. FRIENDLINESS.** Is he a likeable person? Will his fellow-workers and subordinates be drawn to him, or kept at a distance? Does he command personal loyalty and devotion?
- 9. PERSONAL FITNESS FOR THE POSITION.** In the light of all the evidence regarding this applicant, would you recommend him for such a position? Recalling that it is not in his best interest to recommend him for such a position if he is better suited for something else, would you urge him to undertake this work? Do you endorse his application?

Fuller instructions and space for comments on applicant's behavior will be found on the back of this sheet.

Date.....

Applicant's Name or Identification Number.....  
Kind of work for which his suitability is appraised.....

Irritating or indistinct	Understandable or unpleasant	Neither pleasant nor unpleasant	Definitely pleasant and distinct	Exceptionally clear and pleasing
Unimpressive or unimpressive	Creates rather poor impression	Suitable or acceptable	Creates distinctly favorable impression	Impressive or admirable
Slow in grasping meaning of questions	Slow to understand	Nearly always understands questions	Earlier quick in grasping meaning of questions	Exceptionally quick to understand
Confused and illogical	Tends to wander or to become involved	Usually gets his ideas across well	Shows superior ability in clear and convincing	Unusually logical and convincing
Notably lacking in tact and restraint	Shows some tendency to speak impudently and without restraint	Acts judiciously in circumstances requiring tact and judgment	Shows tact and judgment in all circumstances	Impressive evidence of probable soundness of judgment
Over-sensitive or easily discouraged	Occasionally impatient or irritated	Well poised most of the time	Superior self-command	Shows superior self-command and good humor under stress
Timid, hesitant or easily influenced	Appears to be over-conditions	Moderately confident of himself	Wholesomely self-confident	Shows superior self-command
Keeps people at a distance	Does not easily attract friends	Approachable and likeable	Draws many friends to him	An inspirer of personal loyalty and loyalty
Limited for this position	Might do well in some other line of work. Not endorsed	Endorsed	Endorsed with confidence	Endorsed with enthusiasm

SIGNATURE OF RATER

This rating form prepared from suggestions furnished by W. V. Bingham.  
Copyright, 1938, by the International Business Machines Corporation.  
Pat. Form No. 11-200 A 198  
Printed in U. S. A.

FIGURE 2. RATING FORM (Developed and Copyrighted by International Business Machines Corporation)



# INSTRUCTIONS TO INTERVIEWERS AND ORAL EXAMINERS

You are to rate the candidate on certain characteristics which have a bearing on the likelihood that he will be successful in the position for which he is an applicant but which are not measured by a rating of his experience and training, nor by his performance in a written examination, but which can be observed when you talk with him face to face.

Keeping in mind the kinds of duties the candidate will be called upon to perform, consider whether his personal characteristics, as they reveal themselves during the interview, will be an asset or a liability in filling such a position. Do not rate him on his technical knowledge or lack of knowledge of the job. Do not let your estimates of his personal qualities be colored by what he may tell you about his experience or lack of experience.

If his voice, for example, is so rasping or weak that it would give to those with whom he talks an unfavorable impression, you will rate him low in this trait, toward the left end of that scale. If it is neither noticeably pleasant nor unpleasant you should rate him at or near the middle of the scale. If his speech is free from disturbing peculiarities of accent and his voice so clear and resonant that it would be a distinct asset in the work he will do, rate him somewhere on the right half of the scale.

Similarly, rate the candidate on each of the other traits, keeping in mind the definitions of these traits as given on the Rating Form. If a candidate has made no impression on you whatever, either favorable or unfavorable, so far as one of these traits is concerned, rate him at the midpoint of that particular scale. Record your tentative rating on each trait, by putting a check-mark (✓) on the proper scale at the point where, in your judgment, the candidate belongs.

A rating need not fall exactly at one of the subdivisions of the scale. You may place it at either extreme, or anywhere between.

Base your estimates of the applicant's characteristics solely on evidences observed during the interview.

Be sure to record your rating of the applicant on each of the traits. Do not omit any.

When rating the last trait, "Personal Fitness," if—quite apart from any inexperience or lack of technical knowledge which he may have revealed—you consider him definitely unsuited for the position, rate him far to the left. If he barely qualifies, that is, if you are prepared to endorse him as personally suited for this work but can endorse him only with some hesitance, mark him midway between the lower end and the middle of the scale. If you can endorse him with confidence, or with enthusiasm, place your check mark well to the right of the middle.

After you have interviewed several candidates and discussed the evidences of their personal suitability with other examiners, you are at liberty to revise your ratings. This is done by putting your initials above each new check-mark you make.

You may use for comments the appropriate spaces to the right.

## SPACE FOR COMMENT

### 1. VOICE AND SPEECH

### 2. APPEARANCE

### 3. ALERTNESS

### 4. ABILITY TO PRESENT IDEAS

### 5. JUDGMENT

### 6. EMOTIONAL STABILITY

### 7. SELF-CONFIDENCE

### 8. FRIENDLINESS

### 9. PERSONAL FITNESS FOR THE POSITION

### SUGGESTIONS

FIGURE 2. (Reverse)

CITY OF LOS ANGELES  
CIVIL SERVICE COMMISSION  
Personal Qualifications Rating Sheet

EXAMINATION \_\_\_\_\_

Date \_\_\_\_\_ Name of Candidate \_\_\_\_\_

1. APPEARANCE: Is his appearance pleasing? Is he dressed neatly and appropriately? Or is he negligent or slovenly in dress or person? Do you receive the impression that he is habitually careful of his person and dress?  
8   9   10   11   12
  
2. EFFECTIVENESS OF EXPRESSION: Does the candidate express himself well and clearly? In presenting ideas, does he speak appropriately and convincingly? Is his speech marred by defects such as stammering, talking too fast, or affectation? Is his English good?  
8   9   10   11   12
  
3. BEARING AND MANNER: Does the candidate seem assured and confident without being flippant? Is he positive and forceful in his responses without being arbitrary or dogmatic? Or is he excessively nervous and uncertain of himself? Does his manner inspire confidence?  
8   9   10   11   12
  
4. ALERTNESS: Does he seem to grasp the meaning of questions and understand new ideas quickly even when they are complex and difficult? Or does he seem slow in comprehending, confused by details, or does he lose sight of the main points? Does his thinking appear to be clear and logical?  
8   9   10   11   12
  
5. ABILITY TO GET ALONG WITH OTHERS: Can he hold an opinion or point of view with assurance, courteousness, and a friendly spirit of cooperativeness without giving or taking offense? Or does he seem antagonistic or tactless and undiplomatic? Does his attitude and bearing seem to be such as to stimulate antagonism?  
8   9   10   11   12
  
6. MATURITY OF JUDGMENT: Does the candidate impress you as one whose judgment would be dependable on important questions? Does he weigh various sides of problems carefully before arriving at a decision, or does he appear to make snap judgments based only on superficial knowledge? Is his mind closed to new ideas, or does he approach problems with an open mind?  
8   9   10   11   12
  
7. SUMMARY EVALUATION: Does the candidate impress you as being a person who would be thoroughly satisfactory in this position? To what extent would you endorse the candidate? With what assurance would you hire this candidate for this position if you were the appointing officer?  
8   9   10   11   12

Total \_\_\_\_\_

FIGURE 3. RATING FORM  
(Los Angeles City Civil Service Commission)



MS8910

Rating Scale

- A - Excellent  
 B - Very Satisfactory  
 C - Satisfactory  
 D - Passing  
 E - Failing

## MUNICIPAL CIVIL SERVICE COMMISSION

CITY OF \_\_\_\_\_

## QUALIFICATIONS APPRAISAL

POSITION \_\_\_\_\_

Date \_\_\_\_\_

Place \_\_\_\_\_

Candidate No. \_\_\_\_\_

	F	D	C	B	A
<u>EDUCATION</u> Indicate nearest completed year	Minimum required for the position	1 year beyond minimum requirement	2 yrs. beyond minimum requirement	3 yrs. beyond minimum requirement	4 or more yrs. beyond minimum requirement
<u>EXPERIENCE</u>	Minimum required for the position	1 year additional in the field or 2 yrs. additional in a related field	2 yrs. additional in the field or 4 yrs. additional in a related field	3 yrs. additional in the field or 6 yrs. additional in a related field	4 or more yrs. additional in the field or 8 or more yrs. in a related field
<u>VOCATIONAL PROGRESS</u>	No progress; several discharges	Very little advancement; job changes	Some progress	Noticeable progress	Excellent growth and development

## PERSONAL TRAITS

<u>APPEARANCE</u> Physique	Disabling defect	Frail	Acceptable	Energetic	Excellent
Facial Features	Repulsive	Unattractive	Plain	Pleasing	Attractive
Grooming	Offensive	Careless	Plain	Neat	Tasteful
<u>MANNER</u> Approach	Timid or Overbearing	Hesitant or too forward	Modest	Confident	Poised
Diction	Mumbling	Indistinct	Clear	Distinct	Precise
Speech	Laborious	Impulsive	Easy	Accurate	Concise
Voice	Harsh	Irritating	Clear	Pleasant	Well-modulated
Diplomacy	Blundering	Tactless	Cautious	Tactful	Diplomatic
<u>TEMPERAMENT</u> Outlook	Depressed	Easy-going	Hopeful	Cheerful	Enthusiastic
Attitude	Conceited	Self-centered	Satisfied	Unpretending	Objective
Nature	Lazy	Passive	Alert	Enterprising	Vigorous
Interest	In pay only	Indifferent	Average	Keen interest	Eager
<u>SUMMARY</u>	Disqualifying for this work. Not endorsed	Might do well. Endorsed with hesitance	Endorsed	Endorsed with confidence	Endorsed with enthusiasm

SIGNATURE OF INTERVIEWER \_\_\_\_\_

INTERVIEW SCORE \_\_\_\_\_

Use Back of This Sheet for Comments

FIGURE 5. RATING FORM  
 (Municipal Service Bureau, New York State Department of Civil Service)

